DOCUMENT RESUME

ED 108 769 95 PS 008 092

AUTHOR	Cline, Marvin G.; And Others
TITLE	Education as Experimentation: Evaluation of the Follow Through Planned Variation Model. Volume 2B: Monographs and Appendices. Final Report.
INSTITUTION	Abt Associates, Inc. Cambridge, Mass.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Office of Planning, Budgeting, and Evaluation.
PUB DATE	1 May 75
CONTRACT	OEC-0-72-5221
NOTE	312p.; For Volumes 1A and 1B of this report, see ED 094 890 and ED 094 891 for, Volume 2A, see PS 008 091
EDRS PRICE	MF-$0.76 HC-$15.86 PLUS POSTAGE
DESCRIPTORS	Academic Achievement; Community Characteristics; Comparative Statistics; Environmental Influences; Grade 1; Kindergarten; *Longitudinal Studies; Measurement Techniques; *Models; *Primary Education; *Program Evaluation; Research Problems; School Conditions; *Statistical Analysis; Summative Evaluation; Teacher Characteristics
IDENTIFIERS	*Project Follow Through

ABSTRACT

This volume presents monographs relating to the second annual report of the National Evaluation of Follow Through. Monograph 1 is concerned with site-by-site analyses of academic achievement scores for children after two years in the same Follow Through program. Conclusions dealing with first grade vs. kindergarten effects, site effects, and the difficulty in relating findings to specific programs are offered. Monograph 2, which describes the background characteristics and attitudes of the teachers associated with the Follow Through programs, is based on responses to the Spring 1973 Teacher Questionnaire. Monograph 3 focuses on school and community characteristics. Monograph 4 describes Head Start Planned Variation/Follow Through (HSPV/FT) data weaknesses and constraints, and attempts to compare test scores of "treated" and "untreated" children. Appendices and tables containing related data make up approximately one-third of the document. (ED)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED103769

Education as Experimentation: A Planned Variation Model

BLACKBOARD ERASER
PHONICS WORKBOOK
Book B

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: PS TM

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

PS 008092

Abt Associates Inc.

Volume IIB : Appendices and Monographs

ED108769

ABT ASSOCIATES INC.
55 WHEELER STREET, CAMBRIDGE, MASSACHUSETTS 02138
TELEPHONE • AREA 617-492-7100
TELEX: 710-320-6367

Contract No. OEO-0-72-5221

FINAL REPORT

EDUCATION AS EXPERIMENTATION:
EVALUATION OF THE FOLLOW THROUGH
PLANNED VARIATION MODEL

VOLUME IIB:
MONOGRAPHS AND APPENDICES

May 1, 1975

Principal Authors:

Marvin G. Cline
Project Director

N. Ames, R. Anderson, M. Conte, T. Ferb,
C. Hall, M. Joshi, J. Larson, D. Park,
E. Proper, L. Stebbins, C. Stern

Prepared for:

The U. S. OFFICE OF EDUCATION

Office of Planning, Budgeting, and Evaluation
Department of Health, Education, and Welfare
Washington, D.C. 20202

Contract Manager	Richard R. Quality Control Reviewer	Management Reviewer

EDUCATION AS EXPERIMENTATION:

EVALUATION OF THE FOLLOW THROUGH PLANNED VARIATION MODEL

VOLUME IIB

MONOGRAPHS AND APPENDICES

TABLE OF CONTENTS

MONOGRAPH I: AN INTERIM ANALYSIS OF LONGITUDINAL FT/NFT SITE LEVEL DIFFERENCES IN ACADEMIC ACHIEVEMENT

1.0 INTRODUCTION I-1

1.1 Far West Laboratory, Responsive Education Program I-12

  1.1.1 Site Description I-12
  1.1.2 Effects by Site I-13
  1.1.3 Summary I-15

1.2 University of Arizona, Tucson Early Education Model I-21

  1.2.1 Site Descriptions I-21
  1.2.2 Effects by Site I-22
  1.2.3 Summary I-22

1.3 Bank Street College Approach I-28

  1.3.1 Site Descriptions I-28
  1.3.2 Effects by Site I-29
  1.3.3 Summary I-29

1.4 University of Oregon I-32

  1.4.1 Site Descriptions I-32
  1.4.2 Effects by Site I-34
  1.4.3 Summary I-38

1.5 University of Kansas Behavior Analysis Approach I-40

  1.5.1 Site Descriptions I-40
    1.5.1.1 Urban Sites I-40
    1.5.1.2 Rural Sites I-41
  1.5.2 Effects by Site I-42
    1.5.2.1 Urban Sites I-42
    1.5.2.2 Rural Site, Effects of Program Within Site I-46
  1.5.3 Summary I-46

1.6 High Scope Foundation, Cognitively Oriented Curriculum I-49

  1.6.1 Site Descriptions I-49
    1.6.1.1 Closely Matched Sites I-49
    1.6.1.2 Mismatched Sites I-49
  1.6.2 Effects by Site I-49
  1.6.3 Summary I-50

TABLE OF CONTENTS (con'td)

MONOGRAPH I (continued)

1.7 University of Florida Parent Education Model I-58
- 1.7.1 Site Descriptions I-58
- 1.7.2 Effects by Site I-59
- 1.7.3 Summary I-59

1.8 The Educational Development Center (EDC) I-65
- 1.8.1 Site Descriptions I-65
- 1.8.2 Effects by Site I-66
- 1.8.3 Summary I-66

1.9 University of Pittsburgh, Primary Education Project I-71
- 1.9.1 Site Descriptions I-71
- 1.9.2 Effects by Site I-72
- 1.9.3 Summary I-72

1.10 Southwest Education Development Laboratory (SEDL) I-75
- 1.10.1 Site Descriptions I-75
- 1.10.2 Effects by Site I-76
- 1.10.3 Summary I-76

1.11 Conclusions I-80

MONOGRAPH II: TEACHER STUDIES

1.0 INTRODUCTION II-1

2.0 METHOD II-2

2.1 The Sample II-2
2.2 Variable Development II-6
2.3 Analytic Techniques II-14

3.0 RESULTS II-14

3.1 Personal and Professional Characteristics II-14
- 3.1.1 Age and Experience II-14
- 3.1.2 Education and Certification II-20
- 3.1.3 Salary II-23
- 3.1.4 Ethnicity II-25

3.2 Teacher Satisfaction with FT Sponsor II-25
3.3 Amount of Teacher Training II-28
3.4 Reported Teacher Practices II-34
3.5 Parent Involvement II-38
3.6 Teacher Satisfaction with Pulil's Progress II-46

4.0 CONCLUSIONS II-46

TABLE OF CONTENTS (cont'd)

MONOGRAPH III: SCHOOL ENTRY

1.0 INTRODUCTION III-1

2.0 METHOD III-1

2.1 The Sample III-1
2.2 Variables III-8
2.3 Analytic Techniques III-11

3.0 RESULTS III-12

3.1 School Size and Structure III-12
3.2 School Funds III-19
3.3 School Services III-29
3.4 School District Characteristics III-48

4.0 DISCUSSION III-60

MONOGRAPH IV: FOLLOWING THROUGH ON HEAD START VARIATIONS

1.0 THE ORIGINS OF HEAD START PLANNED VARIATIONS IV-1

2.0 THE HSPV/FT STUDIES: INTRODUCTION IV-3

2.1 Participating Groups IV-3
2.2 Background Characteristics IV-7
2.3 Measures IV-11
2.4 The Logic of the Three HSPV/FT Studies IV-13

3.0 STUDY I: KINDERGARTEN AND FIRST GRADE ACADEMIC CORRELATES OF HSPV AND OTHER PRESCHOOL EXPERIENCES IV-14

3.1 The Sample IV-14
3.2 Longitudinal Performance Patterns for Treatment Groups Within the Seven Sites IV-16

4.0 STUDY II: PATTERNS OF SITE MEANS AND RANKS ON ACADEMIC AND AFFECTIVE MEASURES AMONG HSPV/FT CHILDREN IV-25

5.0 CONCLUSIONS IV-43

APPENDICES

APPENDIX A: SCORING PROCEDURES A-1

APPENDIX B: TRANSFORMATIONS B-1

APPENDIX C: RELIABILITIES, VALIDITIES AND COMPARATIVE DATA C-1

- 1.0 Introduction C-1
- 2.0 Reliabilities C-1
- 3.0 Validities C-5
- 4.0 Normative Description C-9

LIST OF TABLES

Page

MII-1 Distribution of Teachers in FT/NFT Studies by Sponsor, Cohort, Grade Level and Region II-3

MII-2 Distribution of Teachers in FT/NFT Studies by Sponsor, Cohort, Grade Level and City Size II-4

MII-3 Distribution of Original Teacher Sample by NFT/FT, Sponsor, Cohort and Grade Level II-5

MII-4 Distribution of Teachers in FT/NFT Studies by Sponsor, Cohort and Grade Level II-7

MII-5 Distribution of Cohort III Non-Entering First Grade Teachers in Studies of Program Related Variables by FT/NFT II-8

MII-6 Selected Cluster Variables Developed From the Spring, 1973 Teacher Questionnaire II-9

MII-7 Predictor Coding Scheme for the Nested Analysis of Teacher Outcomes II-15

MII-8 Means and Standard Deviations for Teacher Age by FT/NFT, Sponsor and Cohort/Grade II-16

MII-9 Means and Standard Deviations for Years of Total Teaching Experience by FT/NFT, Sponsor and Cohort/Grade II-17

MII-10 Means and Standard Deviations for Years of Teaching Experience in Current School by FT/NFT, Sponsor and Cohort/Grade II-18

MII-11 Means and Standard Deviations for Number of Years in Sponsor's Approach by Sponsor and Cohort/Grade II-19

MII-12 Percentage of Teachers with Bachelor's Degrees or Less, More than a Bachelor's Degree but Less than a Master's Degree, or Master's Degree or More by FT/NFT, Sponsor, and Cohort/Grade II-21

MII-13 Percentage of Teachers with Temporary, Renewable, or Permanent Certification by FT/NFT, Sponsor, and Cohort/Grade II-22

MII-14 Means and Standard Deviations for Teacher Salary by FT/NFT, Sponsor and Cohort/Grade II-24

MII-15 Percentage of White Teachers by FT/NFT, Sponsor and Cohort/Grade II-26

MII-16 Means and Standard Deviations for Teachers' Satisfaction by Sponsor and Cohort/Grade II-27

MII-17 Sponsor Profile for Amount of Teacher Training: Social/Emotional/Intellectual Development and Individualization of Instruction II-30

MII-18 Sponsor Profile for Amount of Teacher Training: Classroom Management II-31

LIST OF TABLES (cont'd)

		Page
MII-19	Sponsor Profile for Amount of Teacher Training: Parent Involvement	II-32
MII-20	Sponsor Profiles of Covariate Adjusted Scores in Standard Deviations Units for the Amount of Training Given in Three Areas: Social/Emotional/Intellectual Development, Classroom Management, and Parent Involvement	II-33
MII-21	Sponsor Profile for Reported Teacher Practices: Flexible/Open Classroom	II-35
MII-22	Sponsor Profile for Reported Teacher Practices: Individualization of Instruction	II-36
MII-23	Sponsor Profile for Reported Teacher Practices: Teacher Centeredness	II-37
MII-24	Sponsor Profiles of Covariate Adjusted Scores in Standard Deviation Units for Reported Teacher Practices: Flexible/Open Classroom, Individualization of Instruction, Teacher Centeredness	II-39
MII-25	Sponsor Profile for Importance of Parent Involvement: Parent Decision Making Activities	II-40
MII-26	Sponsor Profile for Importance of Parent Involvement: Housekeeping Activities	II-42
MII-27	Sponsor Profile for Benefit of Parent Involvement	II-43
MII-28	Sponsor Profiles of Covariate Adjusted Scores in Standard Deviation Units for Importance of Parent Involvement: Decision Making Activities, Housekeeping Activities, and Benefit of Parent Involvement	II-44
MII-29	Means and Standard Deviations for Amount of PAC Activity for FT Teachers by Sponsor and Cohort/Grade	II-45
MII-30	Sponsor Profile for Teacher Satisfaction with Pupils' Progress in Reading	II-47
MIII-1	School Questionnaire	III-2
MIII-2	Number of FT and NFT Schools for Each Sponsor	III-9
MIII-3	Frequency Counts of Entering Grade in School for Each Sponsor's FT and NFT Group	III-17
MIII-4	Frequency Counts of Highest Grade in School for Each Sponsor's FT and NFT Group	III-18
MIII-5	Frequency Counts of Total School Enrollment for Each Sponsor	III-20
MIII-6	Frequency Counts of Pupil/Teacher Ratio for Each Sponsor	III-21
MIII-7	Frequency Counts of Receipt of Title I and Other Federal or State Funds for Follow Through and Non-Follow Through	III-22
MIII-8	Frequency Counts of Receipt of Title II and Other Federal or State Funds for Each Sponsor	III-26

LIST OF TABLES (cont'd)

Page

MIII-9 Frequency Counts of Title I Dollars Per Pupil for Each Sponsor III-27

MIII-10 Frequency Counts of Other Federal or State Dollars per Pupil for Each Sponsor III-28

MIII-11 Frequency Distributions of Professional Services for Each Sponsor III-31

MIII-12 Frequency Distributions of Medical Services Personnel for Follow Through and Non-Follow Through III-33

MIII-13 Frequency Distributions of Social Services Personnel for Follow Through and Non-Follow Through III-36

MIII-14 Frequency Distributions of Psychological Services Personnel for Follow Through and Non-Follow Through III-38

MIII-15 Frequency Distributions of Instruction Personnel for Follow Through and Non-Follow Through III-40

MIII-16 Frequency Distributions of Non-Professional Services Personnel for Follow Through and Non-Follow Through III-45

MIII-17 Service Components as a Percentage of Total Services for Follow Through and Non-Follow Through III-46

MIII-18 Frequency Distributions of Professional Services for Each Sponsor III-47

MIII-19 Means and Standard Deviations of Percent Persons in Census Occupation Categories for Sponsors Overall III-49

MIII-20 Means and Standard Deviations of Percent Persons in Census Occupation Categories for Each Sponsor III-50

MIII-21 Means and Standard Deviations of Percent Females in Labor Force for Each Sponsor III-52

MIII-22 Means and Standard Deviations of Average Family Income for Each Sponsor III-53

MIII-23 Means and Standard Deviations of Average Years of Education for Each Sponsor III-54

MIII-24 Means and Standard Deviations of Percent Persons Always Lived in Same House for Each Sponsor III-55

MIII-25 Means and Standard Deviations of Percent Persons Below Poverty Level for Each Sponsor III-56

MIII-26 Means and Standard Deviations of Percent Families with Children Under 18 with Female Heads of Household for Each Sponsor III-58

MIII-27 Means and Standard Deviations of Percent Spanish Speaking and Percent Non-English Speaking Persons for Each Sponsor III-59

		Page
C-1	Reliability Coefficients for a Random Sample of the Population Present for Testing [illegible] Fall 1971	C-2
C-2	Reliability Coefficients for ANCOVA Analytic Sample	C-3
C-3	Interest Correlation Matrix for ANCOVA Analytic Sample	C-6
C-4	Number of Pupils in Random Sample of Population Attaining Scores within Ten Points of Highest Possible Score on WRAT Subtests	C-11
C-5	Comparative Data on WRAT (Level 1): Means and Standard Deviations for Publisher and for ANCOVA Analytic Sample for Ages Five to Ten	C-13
C-6	Comparative Data on WRAT (Level 1): Reliabilities for Publisher's Norming Sample and ANCOVA Analytic Sample	C-15
C-7	Number of Pupils in Random Sample of Population Achieving a Score of 60 or More on the Fall 1971 and Spring 1972 Administration of the 150 Item PPVT (Form A)	C-16
C-8	Comparative Data on PPVT (Form A): Means, Medians, and Standard Deviations for Publisher's Norming Sample and ANCOVA Analytic Sample	C-17
C-9	Comparative Data on MAT (Form F): Means, Medians, and Standard Deviations for Publisher's Norming Sample and ANCOVA Analytic Sample	C-19

LIST OF FIGURES

I-1 First Grade Outcome Profile for Far West Site NC2 I-13
I-2 First Grade Outcome Profile for Far West Site NC1 I-15
I-3 First Grade Outcome Profile for Far West Site W1 I-16
I-4 First Grade Outcome Profile for Far West Site NE I-17
I-5 First Grade Outcome Profile for Far West Site W2 I-18
I-6 First Grade Outcome Profile for Arizona Site S I-22
I-7 First Grade Outcome Profile for Arizona Site NC1 I-23
I-8 First Grade Outcome Profile for Arizona Site NC2 I-24
I-9 First Grade Outcome Profile for Arizona Site NE1 I-25
I-10 First Grade Outcome Profile for Arizona Site NE2 I-26
I-11 First Grade Outcome Profile for Bank Street Site NE1 I-29
I-12 First Grade Outcome Profile for Bank Street Site NE2 I-30
I-13 First Grade Outcome Profile for Oregon Site NC1 I-34
I-14 First Grade Outcome Profile for Oregon Site NC2 I-35
I-15 First Grade Outcome Profile for Oregon Site NE I-36
I-16 First Grade Outcome Profile for Kansas Site NE I-42
I-17 First Grade Outcome Profile for Kansas Site NC2 I-43
I-18 First Grade Outcome Profile for Kansas Site S I-44
I-19 First Grade Outcome Profile for Kansas Site NC1 I-46
I-20 First Grade Outcome Profile for High Scope Site NE I-50
I-21 First Grade Outcome Profile for High Scope Site NC I-51
I-22 First Grade Outcome Profile for High Scope Site W1 I-52
I-23 First Grade Outcome Profile for High Scope Site W2 I-53
I-24 First Grade Outcome Profile for High Scope Site S I-54
I-25 First Grade Outcome Profile for High Scope Site W3 I-55
I-26 First Grade Outcome Profile for Florida Site S1 I-59
I-27 First Grade Outcome Profile for Florida Site S2 I-60
I-28 First Grade Outcome Profile for Florida Site W I-61
I-29 First Grade Outcome Profile for Florida Site NC I-62
I-30 First Grade Outcome Profile for EDC Site NE1 I-66
I-31 First Grade Outcome Profile for EDC Site NE2 I-67
I-32 First Grade Outcome Profile for EDC Site S I-68
I-33 First Grade Outcome Profile for Pittsburgh Site NC I-72

LIST OF FIGURES (cont'd)

		Page
I-34	First Grade Outcome Profile for Pittsburgh Site NE	I-73
I-35	First Grade Outcome Profile for SEDL Site NE	I-76
I-36	First Grade Outcome Profile for SEDL Site W	I-77
I-37	First Grade Outcome Profile for SEDL Site S	I-78

LIST OF FIGURES (cont'd)

		Page
MIII-1	Frequency Distributions of Entry Grade for Follow Through and Non-Follow Through	III-13
MIII-2	Frequency Distribution of Highest Grade in School for Follow Through and Non-Follow Through	III-14
MIII-3	Frequency Distributions of Total School Enrollment for Follow Through and Non-Follow Through	III-15
MIII-4	Frequency Distributions of Pupil-Teacher Ratio for Follow Through and Non-Follow Through	III-16
MIII-5	Percentage of Title I Dollars per Pupil by FT/NFT	III-23
MIII-6	Frequency Distributions of Other Federal or State Funds per Pupil for Follow Through and Non-Follow Through	III-24
MIII-7	Frequency Distributions of Professional Services for Follow Through and Non-Follow Through	III-30
MIII-8	Frequency Distributions of Medical Services for Follow Through and Non-Follow Through	III-32
MIII-9	Frequency Distributions of Social Services for Follow Through and Non-Follow Through	III-35
MIII-10	Frequency Distributions of Psychological Services for Follow Through and Non-Follow Through	III-37
MIII-11	Frequency Distributions of Instructional Services for Follow Through and Non-Follow Through	III-39
MIII-12	Frequency Distributions of Activities for Follow Through and Non-Follow Through	III-41
MIII-13	Frequency Distributions of Library Services for Follow Through and Non-Follow Through	III-43
MIII-14	Frequency Distributions for Non-Professional Services for Follow Through and Non-Follow Through	III-44
MIV-1	Sponsor and Cohort Distribution of Traced Children	IV-4
MIV-2	Treatment History Distribution of HSPV-FT Children in Cohort III	IV-5
MIV-3	Treatment History Distribution of HSPV-FT Children in Cohort IV	IV-6
MIV-4	Background Variables for Traced Children and for Various Tested Subsets of Traced Children	IV-8
MIV-5	Structure of Three-Point Longitudinal Samples	IV-9
MIV-6	Site Means of Background Variables at Four Testing Points, Cohort II Traced Children	IV-10
MIV-7	Background Characteristics and Mean Test Scores of Children by Site and Treatment Group	IV-15
MIV-8a	Treatment Group Means of Math and Reading Scores for Duluth	IV-17

LIST OF FIGURES (cont'd)

		Page
MIV-8b	Treatment Group Means of Math and Reading Scores for Tacoma	IV-18
MIV-8c	Treatment Group Means of Math and Reading Scores for Lincoln	IV-19
MIV-8d	Treatment Group Means of Math and Reading Scores for Portageville	IV-20
MIV-8e	Treatment Group Means of Math and Reading Scores for Fort Walton Beach	IV-21
MIV-8f	Treatment Group Means of Math and Reading Scores for Greenley	IV-22
MIV-8g	Treatment Group Means of Math and Reading Scores for Seattle	IV-23
MIV-9	Longitudinal Performance Patterns for FT Groups Within Sites	IV-24
MIV-10	Site Means of Background Variables for HSPV-FT Children Tested in Fall of Head Start and in Spring of Grade 1	IV-26
MIV-11a	Site Means and Ranks of Means of Pretest Variables for Children Traced from Pretest (Head Start) to Postest (Spring, Grade 1)	IV-27
MIV-11b	Site Means and Ranks of Means of Fall Kindergarten Test Scores for Children Traced from Pretest to Postest	IV-28
MIV-11c	Site Means and Ranks of Means of Spring Kindergarten Test Scores for Children Traced from Pretest to Postest	IV-29
MIV-11d	Site Means and Ranks of Spring First Grade Postest Scores for Children Traced from Pretest to Postest	IV-30
MIV-12	Aggregate Site Ranks for Five Categories of Tests	IV-31
MIV-13	Longitudinal Patterns of Mean Site Ranks of Reading Scores over Four Testing Occasions	IV-32
MIV-14a	Longitudinal Patterns of Mean Site Ranks of Reading Scores over Four Testing Occasions	IV-35
MIV-14b	Longitudinal Patterns of Ranks of Site Mean Math Scores over Four Testing Occasions	IV-36
MIV-14c	Longitudinal Patterns of Mean Site Ranks of Vocabulary Scores over Four Testing Occasions	IV-37
MIV-15	Longitudinal Patterns of Site Mean Scores on Reading, Math, and Vocabulary Measures, for Seven Sites	IV-39
MIV-16	Sites Showing Gains in Rank Standing on Three Dimensions of Progress During Three Years in HSPV-FT	IV-42

LIST OF FIGURES (cont'd)

		Page
B-1	Raw Scores of Kindergarten Days Absent: Pooled NFT	B-4
B-2	Transformed Scores of Kindergarten Days Absent: Pooled NFT	B-5
B-3	Raw Scores of First Grade Days Absent: Pooled NFT	B-6
B-4	Transformed Scores of First Grade Days Absent: Pooled NFT	B-7
B-5	Raw Scores of Kindergarten Gumpgookies Total Score: Pooled NFT	B-8
B-6	Transformed Scores of the Kindergarten Gumpgookies Total Score: Pooled NFT	B-9
B-7	Raw Score of Kindergarten Gumpgookies Instrumental Activities Subtest: Pooled NFT	B-10
B-8	Transformed Scores of Kindergarten Gumpgookies Instrumental Activities Subtest: Pooled NFT	B-11
B-9	Transformed Scores of Kindergarten Gumpgookies School Enjoyment Subtest: Pooled NFT	B-12
B-10	Raw Scores of Kindergarten Gumpgookies Self-Evaluation Subtest: Pooled NFT	B-13
B-11	Raw Scores of Kindergarten Gumpgookies Self-Confidence Subtest: Pooled NFT	B-14
B-12	Raw Scores of Kindergarten Gumpgookies Purposive Behavior Subtest: Pooled NFT	B-15
B-13	Raw Scores of First Grade MAT Work Knowledge Subtest: Pooled NFT	B-16
B-14	Raw Scores of First Grade MAT Word Analysis Subtest: Pooled NFT	B-17
B-15	Transformed Scores of First Grade MAT Word Analysis Subtest: Pooled NFT	B-18
B-16	Raw Scores of Kindergarten Gumpgookies Total Score: Random Sample of Population	B-19
B-17	Raw Scores of Kindergarten Gumpgookies Instrumental Activities Subtest: Random Sample of Population	B-20
B-18	Raw Scores of Kindergarten Gumpgookies School Enjoyment Subtest: Random Sample of Population	B-21
B-19	Raw Scores of Kindergarten Gumpgookies Self-Evaluation Subtest: Random Sample of Population	B-22
B-20	Raw Scores of Kindergarten Gumpgookies Self-Confidence Subtest: Random Sample of Population	B-23

LIST OF FIGURES (cont'd)

		Page
B-21	Raw Scores of Kindergarten Gumpgookies Purposive Behavior Subtest: Random Sample of Population	B-24
B-22	Raw Scores of First Grade MAT Word Knowledge Subtest: Random Sample of Population	B-25
B-23	Raw Scores of Fist Grade MAT Word Analysis Subtest: Random Sample of Population	B-26

MONOGRAPH I:

AN INTERIM ANALYSIS OF LONGITUDINAL FT/NFT SITE LEVEL DIFFERENCES IN ACADEMIC ACHIEVEMENT

Monograph I

An Interim Analysis of Longitudinal FT/NFT Site Level Differences in Academic Achievement

1.0 Introduction

The data for the children of Cohort III who entered kindergarten in the Fall of 1971 cover participation in the Follow Through experiment for a full two years at the time of this analysis. Many of these children have been repeatedly tested during this interval (at entry in the fall of kindergarten, at exit from kindergarten and at exit from first grade). Thus, this group of children offers a unique opportunity to study not only the status of sponsor's FT and NFT groups at the end of first grade, after two years of school experience, but also the course of events in kindergarten and first grade that has led up to their status at the end of first grade. Since attrition has already substantially reduced the number of children available for analysis (a trend which will no doubt continue). The possibility of a similar analysis using a large group of traced children is somewhat unlikely. The central questions of interest in this study are:

What are the differences between gains in FT and NFT groups over the kindergarten and first grade period?

- which are attributable to kindergarten experience?
- which are attributable to first grade experience?
- which are attributable to differences between FT and NFT at entry into kindergarten?

Do these differences between gains over the two year period suggest any trends?

Do children, FT or NFT, who start with a disadvantage at entry catch up to and/or surpass their comparison group? Do they

- maintain their entry status with respect to their comparison?
- fall progressively further behind their comparison?

Do sponsor's FT groups show consistent/inconsistent patterns across grades? Across sites?

Of special relevance is the fact that the analysis deals with the sites. Here, analyses of the traced sample organized by individual site are reported for each Sponsor. A series of analyses are reported in which the FT children meeting the inclusion criteria within a given site will be contrasted with the NFT children meeting the criteria in the same site.

It is appropriate to indicate here the reasons for the addition of the site analyses to the present chapter. There are three reasons for these additional analyses:

The first reason is that there is a good deal of anecdotal evidence that very divergent socioeducational events are occurring across sites for a given Sponsor. These include the issues of implementation (discussed in Monograph III, Volume IB of this series) which appear to be operating in somewhat different ways in different sites for some Sponsors. The anecdotal evidence suggests that sites are very different in respect to the administration and operation of FT programs and in the characteristics of the comparison NFT groups. Combining sites means combining diverse events, which tends to obscure important processes.

The second reason for the site analyses is that the discrepancies between some of the kindergarten effects found with the traced two-year sample and those found for the one-year kindergarten sample (discussed in the previous chapter) appear in some cases to be related to site differences in the two samples. That is, in some cases the loss of children in the traced group was not evenly distributed over all sites but was concentrated in some sites, and the change in effects between the two sets of analyses appears to be associated with the loss of whole sites and/or disproportionate numbers of children from a few sites. Looking at individual sites is thus justified on these grounds as well as on anecdotal grounds.

The third reason for the site analyses is that a substantial amount of heterogeneity of regression occurs in the overall Sponsor analyses. This means that for a given Sponsor on a given outcome measure, the covariable (i.e., the pretest) performs its adjusting function in one way for the FT group and in another way for the NFT group, and it therefore removes non-comparable portions of the variance from these two groups. Adjustments for differences in pretest may be spuriously high in one group and spuriously low in the other when heterogeneous regressions are ignored, and comparisons between the groups can be gravely misleading.

The critical factor here is that heterogeneity of regression may be produced in overall Sponsor analyses precisely because all sites for a given Sponsor are combined in that Sponsor's analyses. The heterogeneity exists because each site shows its own distinctive relationships between pretest and outcomes. Just as the anecdotal evidence suggests that sites often contain very different educational dynamics, the heterogeneity findings suggest that sites are quite different and should often not be combined in analyses. That this is most likely the case is demonstrated in the elimination of heterogeneity when analyses are performed on each site. When taken one at a time, the sites are homogenous and can be analyzed in the ANCOVA mode.

Site analyses also represent a form of replication of findings to which we have referred earlier as an aspect of the stability of Sponsor effects from place to place. It would be difficult to make any policy decisions about any program if one did not know how likely the observed effects would be to emerge in new sites. The sites in this analysis were not selected by any sampling plan, and so there is no formal statistical justification for any generalization of findings. Nevertheless, it is reasonable to look for stability or instability of findings over several different locations, simply to get a first tentative estimate of the sensitivity of the various FT programs to different conditions and locations of operations. For this reason as well, we report individual site analyses.

Before considering the method to be utilized in this analysis, it is important that the reader be made aware of the exact nature of the analysis. We are concerned in this analysis with differences between particular FT and NFT groups; differences in post test means and differences in gains[1] over various time periods.

An example will help clarify this point.

A given group may show an average gain of 5 points which is attributeable to first grade experience on a first grade outcome, while their comparison group may show an average gain of 10 points. These gains can be arrived at by adjusting an end of first grade outcome measure for differences between the groups on measures prior to first grade, including end of kindergarten achievement tests differences as well as differences on background characteristics and achievement tests administered at entry to kindergarten. This adjustment gives an estimate of gain attributable to first grade experience.

---

[1] Residual gains

Since the present analysis is concerned with differences between gains as an indicator of advantageous educational investment, the absolute value of the gain is not considered and only the differences between gains (in the example given a value of 5 points) are utilized. It is immediately obvious that a given difference of 5 points can be arrived at by subtracting any pair of gains which differ by 5 points, e.g., 0 and 5, 10 and 15, etc. Clearly, differences between gains can have various meanings depending upon the level of the gains. The difference between 10 and 15 points may not have the same educational significance as a difference between 25 and 30 points. We choose, however, for several reasons, to focus this additional longitudinal analysis on the magnitude of differences and not on the level at which the differences occur.

The first reason is that we are interested in the shape of the longitudinal curve since it does reflect the relationship of each position on the curve to subsequent positions. This is the real goal of a longitudinal analysis -- the examination of a pattern over time. The level of the pattern is important but not the same as the question of the sequence of effects over time.

The second reason is that levels differences really represent site differences and are sharply confounded by site-associated events. Thus, a difference between 25 and 30 points indicates that a Sponsor is dealing with a community of high gaining children who are assigned to the FT and NFT groups by the school administration, whereas a difference between 10 and 15 points indicates a community of lower gaining children. A community in which the average gain across all children regardless of what educational experiences they are having lies somewhere between 25 and 30 points is very different from a community in which the average gain for all children falls somewhere between 10 and 15 points. Thus, the levels of the longitudinal data presented here should be understood as the context for interpreting the findings, and not as a factor in the analysis. Each Sponsor will be analyzed by site so that such levels differences will not confound the analysis.

Finally, all differences between FT and NFT groups at each site are transformed into standard scores based upon the standard deviation of the distribution of the relevant outcome measure. Therefore, the differences in gains observed for the several groups are comparable regardless of where in the distribution they have occurred.

These procedures allow us to consider the pattern of effects over time in such a way that the portion of the differences in gain at each time point attributable to Follow Through can be identified. This is the policy question of importance here: to what extent has Follow Through contributed to the ongoing pattern of growth as that pattern is uniquely defined in each site.

Method

A detailed description of the Sponsor's FT and NFT groups examined in this analysis is presented in Appendix A, Vol. IIA, as are various other details of the analysis. The demographic characteristics and pupil entry scores of FT and NFT children in the sites are also discussed in detail in Vol. IIA, Chapter III, and are repeated here. The only unique aspect of the current analysis is the method.

The reader who is not concerned with the details of the method should proceed to the results, after considering the manner of presentation of the results.

The results are presented in charts which show differences between FT and NFT; FT-favoring differences are shown by upward sloping lines, and NFT-favoring differences by downward sloping lines. The figures do not show gains or growth. Rather they show differences in gains between FT and NFT as these differences are related to subsequent differences in such gains.

Figure I-1 serves as an illustrative example. The figure shows the differences between FT and NFT at entry to kindergarten, at exit from kindergarten, and at exit from first grade in standard deviation units of the outcome measures. The distance between the points is intended to reflect the time interval between testing: approximately 3/4 of a year for the kindergarten entry to exit interval and approximately 1 year for the kindergarten exit-first grade exit interval. The meaning of the value at point F (exit from first grade) is the most apparent: the value for MAT Reading, +.70SDs, indicates that the FT group is substantially ahead of the NFT group in absolute terms, since this value represents the standardized unadjusted difference between FT and NFT on the MAT Reading subtest, a difference of 10.01 points.

The value at point K, exit from kindergarten, .44SDs, is computed by subtracting the difference between first grade gain[2] in FT and NFT, +3.61, from the unadjusted difference, 10.01, and converting to standard deviation units. Since the difference between FT and NFT gain during first grade is +3.61, showing that FT gained more than NFT during first grade, the line from K to F has a positive slope, FT-favoring. Furthermore, the magnitude of change from point K, the end of kindergarten, to point F, the end of first grade, shows the difference between FT and NFT gain during first grade in standard deviation units, 3.61/14.26. Point K, marking the end of kindergarten, also serves an additional purpose. By subtracting the first grade difference between FT and NFT gain from the unadjusted difference, the difference between the groups at the end of kindergarten is located. Point K thus shows the difference between FT and NFT at the end of kindergarten, which contributed to the difference at the end of first grade. The value for MAT Reading is .44 standard deviations indicating that the FT children were substantially ahead of NFT at the end of kindergarten, at least on the spring kindergarten WRAT which was utilized in these analyses, and that the end-of-kindergarten difference contributed substantially to the end-of-first-grade differences.

Point P shows the difference between FT and NFT at entry into kindergarten expressed in standard deviation units of the end of first grade outcome. Point P is computed by subtracting the difference between FT and NFT gain over the two year period from entry to kindergarten to exit from first grade, 10.67 from the unadjusted difference, 10.01, and by converting to standard deviation units. This subtraction yields a small negative value, -.04, indicating that the FT and NFT groups were closely matched at entry, on both entry WRAT scores and/or level of mother's education. The line from P to K, entry to kindergarten to exit from kindergarten, shows a positive slope, that is, a FT favoring difference, and reflects the magnitude of the difference between FT and NFT gain in kindergarten which is related to end of first grade differences.

---

[2] The difference between FT and NFT first grade gain is computed by determining the part of the unadjusted difference that is uniquely attributable to first grade experience. This process is fully described in the description of the analytic method.

Variation in the outcome measures and in their pattern of differences across the two year period are shown in the figures as well. Figure I-1 shows that the FT group is pulling ahead in word analysis at a constant rate, that the FT children gained more than the NFT in Reading in both years, but gained relatively more in kindergarten than in first grade, and that the FT group gained substantially more than NFT in Mathematics during kindergarten, but gained much less in first grade. Before turning to the method, a look at Figure I-3 will complete the picture. This figure shows that the FT group started out behind NFT and that the FT group is falling progressively further behind NFT.

Although these figures present longitudinal patterns of difference across kindergarten and first grade, they do not represent achievement domains measured across time. The outcome measures, the MAT Reading, Mathematics, and Word Analysis subtests, do represent achievement domains. However, for the computation of entry and kindergarten components, the fall and spring kindergarten WRAT total scores were utilized. Thus, a given group may show an advantage on kindergarten WRAT subtests, e.g., WRAT Reading or Arithmetic which is not reflected in the longitudinal pattern. This inconsistency only occurs when a kindergarten WRAT subtest difference is not reflected in the total test, an event which occurs when subtests are added and the difference on one subtest is balanced out by differences in the opposite direction on the other subtests.

This inconsistency represents a logical difficulty of the longitudinal patterns. However, since skills and knowledge measured at the end of first grade are a development of many aspects of a child's prior experience, certainly more than that measured by a single subtest of the WRAT, the total test is more appropriate. Furthermore, the correlation between the total WRAT and the outcome measures is somewhat higher than the correlations of WRAT subtests with outcome measures.

For the sake of discussion, it is possible to label Sponsors according to their effects pattern in each outcome measure. A Sponsor can be called an "accumulator" if the two-year effect is the accumulative result of relatively larger NFT gains in both grades. It is likely that first grade gains build upon kindergarten gains in the sense that children who improved the most in kindergarten continued to build their advantage over the NFT in first grade. This is a likely interpretation; however, a definitive examination of this point requires an exploration of gains achieved by different types of children.

A sponsor can be a "kindergartner" because all of the difference on end of first grade outcomes can be attributed to the kindergarten year. A kindergartner may show losses in the first grade, or may show a maintenance of the same relative position in first grade which was achieved in kindergarten. In this case, the "kindergartner" can be termed a first grade "maintainer."

A Sponsor can be a "first grader" because all of the two-year effects are found in the first grade and not in kindergarten. Such a Sponsor may have laid some foundations for first grade growth during the kindergarten year, but these effects did not show up on the achievement battery used during kindergarten. It may also be that such a Sponsor may have an approach which is not appropriate to the educational needs of the kindergartners with whom the Sponsor is working, but is appropriate to the first graders in the sample. A Sponsor who is a kindergartner and not, therefore, a first grader can also be assumed to be laying a foundation in kindergarten, but this time for a weaker rate of gain in first grade. This may be related to the inappropriateness of a program for first grade, or because the first grade test battery is not related to the areas which a program emphasizes. A finer grained analysis than that presented in this Report is required to examine this issue.

Finally, a Sponsor may show negative slopes across grades in which case the children working with this Sponsor gain less than the NFT groups during both grades. It seems unlikely that any of the educational models utilized in the planned variation experiment can be expected to retard the growth of children vis-a-vis NFT children when the groups are matched. The more likely explanation of such a pattern is that the program emphasis of such Sponsors are those which are not picked up by the traditional achievement tests used in the upper grades, despite their absence in the kindergarten and first grade tests. On the other hand, an accumulation of negative effects may be the result of very special circumstances at the point of administration of the model which attenuates or inhibits FT gains relative to the NFT children. Once again, a great deal more data are required to make sense out of a pattern which shows an accumulation of negative effects over two years. Such issues cannot be addressed in this Report.

Three ANOVA/ANCOVA models and their interrelation are used in these analyses:

MODEL I: $Y_F = b_o + b_1T$

where Y_F = expected value on a first grade outcome

b_o = mean of Y_F

b_1 = difference between FT and NFT in units of Y_F

T = a group membership variable coded +.5 for FT membership and -.5 for NFT membership

Model I gives the difference between the FT and NFT outcome means in absolute terms, the unadjusted difference.

MODEL II: $Y_F = b_{ADJ1} + b_2T + b_EX_E$[3]

b_{ADJ1} = mean of Y_F adjusted for kindergarten entry test

b_2 = difference between FT and NFT in units of Y_F adjusted for kindergarten entry test, an estimate of the FT/NFT difference in two year gain.

X_E = kindergarten entry test

b_E = relationship between X_E and Y_F in units of Y_F

Model II gives the difference between FT and NFT on the first grade outcome adjusted for differences on X_E, the entry kindergarten test. The value of the adjustment $[b_E(\overline{X}_{FT} - \overline{X}_{NFT})] = [b_1 - b_2]$ is related to the difference between the FT and NFT group means on the kindergarten entry test. The sign of the adjustment depends on the relation of the FT and NFT means. The adjustment $[b_1 - b_2]$ is thus the part of the difference between FT and NFT on the first grade outcome that is attributable to entry test differences. This adjustment can also be viewed as the expected difference between FT and NFT, partialing out both kindergarten and first grade experience.

---

[3] This is the model used to estimate adjusted effects in Volume IIA.

MODEL III: $Y_F = b_{ADJ2} + b_3T + b_EX_E + b_KX_K$

where b_{ADJ2} = mean of Y_F adjusted for entry kindergarten (b_EX_E) and exit kindergarten (b_KX_K) differences between FT and NFT

b_3 = the differences between FT and NFT similarly adjusted, an estimate of the FT/NFT difference in first grade gain

b_E = relationship between X_E, kindergarten entry test, and Y_F, the first grade outcome adjusted for the kindergarten exit outcome.

X_K = the kindergarten exit outcome

b_K = the relationship between Y_K and Y_F, the first grade outcome, adjusted for the kindergarten entry outcome

Model III gives the difference between FT and NFT on the first grade outcome adjusted for both kindergarten entry and exit differences. The difference, b_3, is the part of the difference on the first grade outcome that is attributable to first grade experience.

Again, the adjustment is of interest. The difference between the unadjusted effect, b_1, and the adjusted effect, b_3, defines the outcome attributable to experience prior to first grade--the combination of background experience as measured by kindergarten entry tests and kindergarten experience as measured by kindergarten exit tests.

The models thus give estimates of:

- the expected difference on the first grade outcome based on differences at kindergarten entry, P, $P = b_1 - b_2$

- the expected difference on the first grade outcome based on the combined differences on kindergarten entry and exit tests, K, $K = b_1 - b_3$
- the unadjusted difference, F $F = b_1$

The results of these analyses are presented in Appendix tables.

1.1 Far West Laboratory, Responsive Educational Program

The Sponsor describes the Responsive Educational Program as one that addresses the variety of learning needs of children while responding to the variety of needs of parents, schools and communities. Responsive teaching focuses on a child's language development and problem solving skills to produce an independent learner. An independent learner has the skills to solve a variety of problems and the confidence to attempt them. The responsive classroom encourages exploration and discovery learning; the child's life experiences are integrated into classroom activities.

All five sites used for the analysis of the Far West Program are metropolitan communities of 100,000 persons or more. However, the sites are scattered geographically and the children in the programs differ greatly in ethnic background, language, SES, and initial achievement.

1.1.1 Site Description

Tables VI-1 through VI-5 in Vol. IIA[4] present pupil and family background characteristics for each of the five sites included in these analyses. The first site (NE) is a large Northeastern city, serving predominantly Black lower SES children (FT = 100%; NFT = 71%). The two groups are similar in kindergarten entry level.

The second site, a medium-sized North Central city (NC1), differs from the first not only in size and region, but in at least two of the characteristics of the children served. The FT/NFT children are predominantly White (FT = 92%; NFT = 93%). They differ substantially in SES; the FT children come from poorer families, characterized by a lower level of parental education and slightly fewer intact marriages. They are, nevertheless, well matched with the NFT group at kindergarten level.

At the third site, a large North Central city (NC2), both FT and NFT groups are virtually all Black and have similar SES, although there is a smaller number of intact families in the FT group. In addition, although the two groups are well matched on the WRAT (FT = 27.46; NFT = 28.80), the FT children enter kindergarten behind their NFT counterparts on the PPVT (FT = 36.63; NFT = 42.26).

At the fourth site, a medium-sized Western city (W1), approximately two-thirds of the children are Black (FT = 60%; NFT = 70%) and one-third are Spanish-surnamed (FT = 35%; NFT = 24%). The FT group has more intact families and slightly higher family incomes. The FT children enter kindergarten well behind the NFT children on the WRAT (FT = 22.15; NFT = 26.82) despite their slightly higher SES and roughly equivalent scores on the PPVT (FT = 44.09; NFT = 45.37). (The difference on the WRAT may reflect the fact that the FT group contains a somewhat higher percentage of Spanish-speaking children; that a corresponding difference is not found on the PPVT may reflect the relative SES of the two groups and/or the benefit of a language-oriented preschool program. A majority of both groups attended preschool.)

---

[4] All Tables for Pupil and Family Background Characteristics are located in Vol. IIA - Two Year Effects of Follow Through.

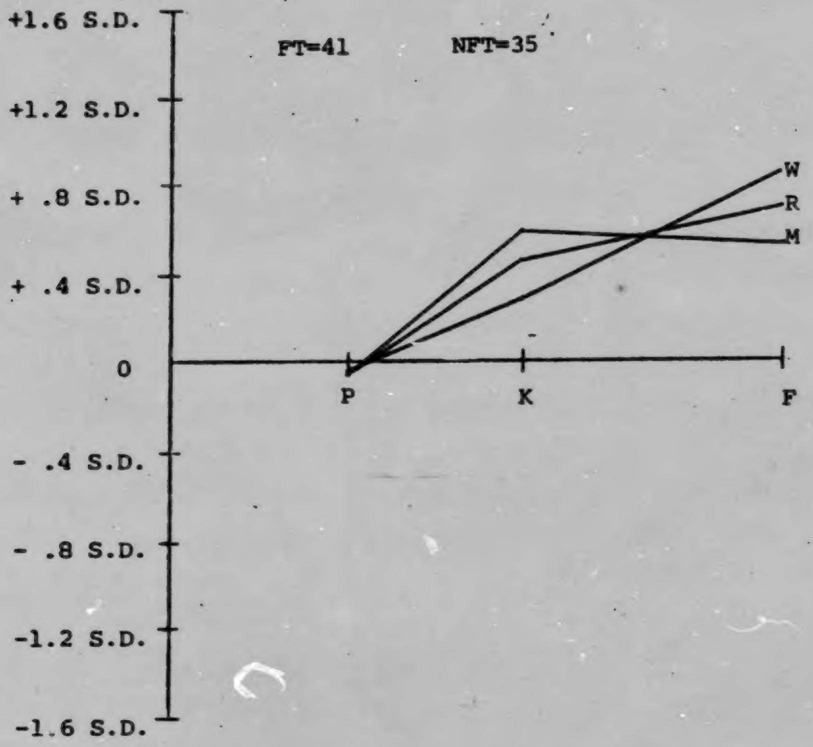
At the last site, another medium-sized Western city (W2), the FT/NFT children are primarily White (FT = 63%; NFT = 76%), although the groups contain Spanish-surnamed children (FT = 22%; NFT = 3%) as well. In this community, the FT children have a substantially lower SES than the NFT children and are slightly behind at kindergarten entry level.

1.1.2 Effects by Site

There appears to be a variety of patterns emerging in the several sites working with this Sponsor as demonstrated in Tables I-1 to I-5. A summary of these various patterns will make this finding obvious.

A clear pattern of strong gains on all outcome measures is demonstrated in site NCs (Table I-1). These gains cumulate over the two years in the Word Analysis measure, and, to a lesser extent, in the Reading test. The FT group's early gains in Math are maintained at the same relatively higher performance level at the end of first grade; that is, the FT children appear to be gaining in Math at a greater rate than the NFT children during the first year of the program. However, the rate of gain appears to be equalized during the second year so that the FT children simply maintain the same relatively high performance level in the first grade.

The pattern of gains in the second north central site (NC1, Table I-2) differs in some respects from the first pattern observed in NC2. Here, there is no differential FT/NFT gains during the first year but the Word Analysis and Reading tests show much greater FT gains by the end of the second year. This is not true in the Math domain where the FT and NFT show essentially the same levels and rates of gains over the two year period. The only obvious similarity between the findings of these two sites is in the Word



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.1785*	5.8418*	10.0167*
S.E.	.0437	2.3998	3.0933
Two Year Difference (B)	.1878*	6.3314*	10.6752*
S.E.	.0395	2.2277	2.7826
First Grade Difference (C)	.1167*	- .5450	3.6182
S.E.	.0477	2.4502	3.2118
Outcome S.D.	.2087	10.7608	14.2659
P=(A-B)/S.D.	- .0446	- .0455	- .0462
K=(A-C)/S.D.	.2961	.5935	.4485
F=A/S.D.	.8553	.5429	.7021

Figure I-1 First Grade Outcome Profile for Far West Site NC2

Note: All items marked by * are significant ($p <$ .05).

Analysis and Reading Gains during the second year of participation in the program. Since these two sites are quite different from each other on several demographic characteristics (SES, rural/urban, ethnic composition, etc.), the similarity of pattern cannot plausibly be attributed to such factors. In order to assess other plausible factors which might be associated with this Sponsor's findings, it is necessary to turn to the findings for the remaining sites.

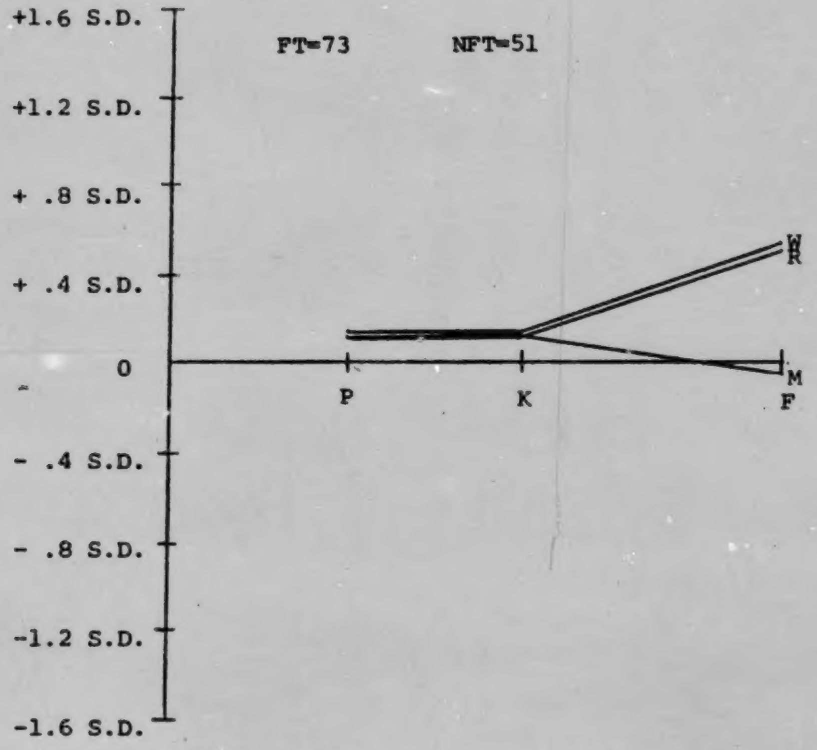
Both the northeastern and the predominantly white western site (NE and W2 respectively) show similar and close to null effects. There are no differences between the FT and NFT groups on Word Analysis and Reading in the W2 site (the two groups are gaining at the same rate throughout the two year span), and there is a slightly greater NFT gain in Math. This same Math differential is present in the NE site, along with a slight NFT advantage in Word Analysis. The two groups are equivalent in Reading.

The last site (W1) shows a pattern that is a somewhat more exaggerated version of the NE and W2 patterns. That is, there is no FT/NFT differences during the first year, but a greater NFT gain in all three domains at the end of the second year. The shape of all three of these patterns is similar in the sense that the entry differences between the two groups of children at each site is maintained during the first year, and there is a tendency toward greater NFT gain during the second year.

Once again, the demographic dissimilarity among these three sites cannot easily be used to account for broad similarity among their findings; nor can the demographics be used to account for the differences between this last set of three patterns and the generally positive findings observed in the first set of two sites.

1.1.3 Summary

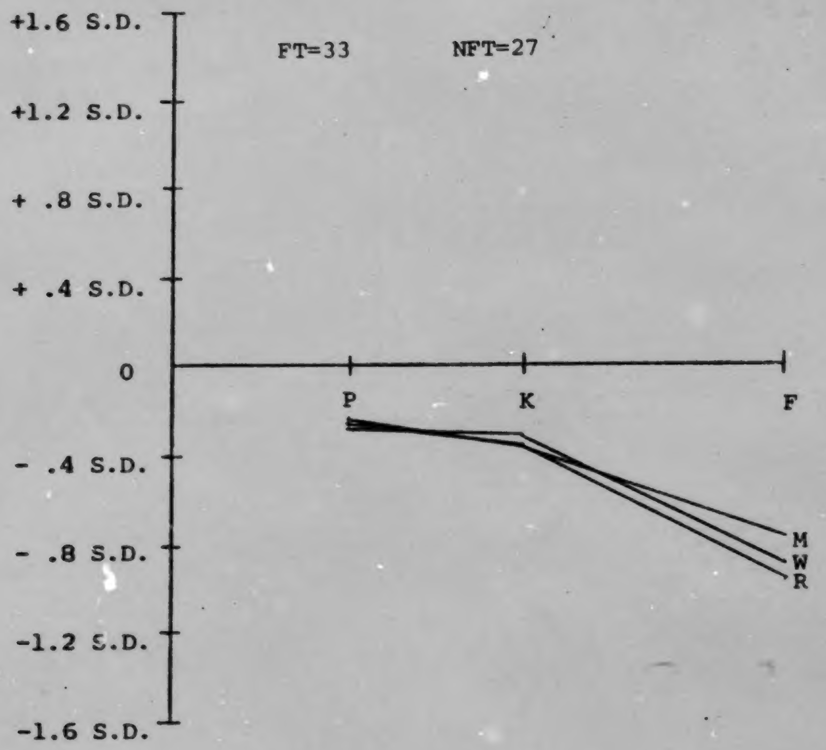
Clearly there are conditions of place, persons, and time under which this program is effective and conditions under which it is not. But these factors are not directly associated with the measures of ethnicity, SES, city size and entry level which differentiate the sites and which constitute the full set of data available for these analyses. At this point in the analyses, there are no data-based explanations for the large amount of variability of effects across sites for the Far West program. It is critical, for the full understanding of this model, and for the assessment of its potentials, to examine the effects



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.1284*	- .5404	8.4735*
S.E.	.0412	1.8307	2.9710
Two Year Difference (B)	.0947*	-1.7556	6.5415*
S.E.	.0305	1.3682	2.1203
First Grade Difference (C)	.0961*	-1.6991	6.6381*
S.E.	.0278	1.2799	1.9506
Outcome S.D.	.2336	9.9938	16.7451
P=(A-B)/S.D.	.1443	.1216	.1154
K=(A-C)/S.D.	.1383	.1159	.1096
F=A/S.D.	.5496	- .0541	.5060

Figure I-2 First Grade Outcome Profile for
Far West Site NC1

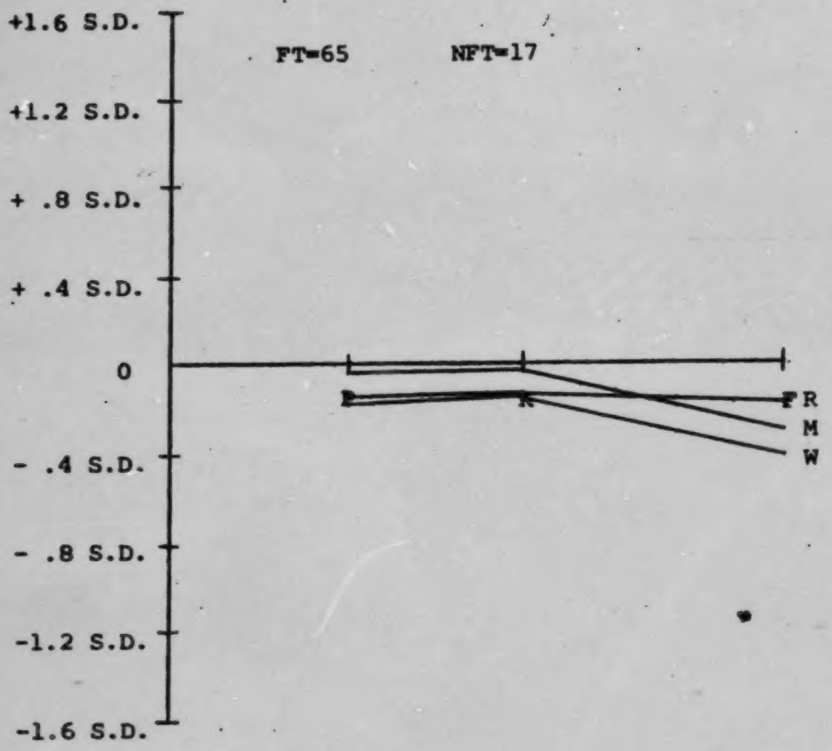
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	-.2122*	- 8.1178*	-11.6397*
S.E.	.0543	2.5317	2.6739
Two Year Difference (B)	-.1459*	- 5.5589*	- 8.6619*
S.E.	.0415	2.0374	2.1615
First Grade Difference (C)	-.1374*	- 4.3776*	- 7.4269*
S.E.	.0431	2.0282	2.1548
Outcome S.D.	.2331	10.4953	11.7677
P=(A-B)/S.D.	-.2844	- .2438	- .2530
K=(A-C)/S.D.	-.3209	- .3564	- .3580
F=A/S.D.	-.9103	- .7735	- .9891

Figure I-3 First Grade Outcome Profile for Far West Site W1

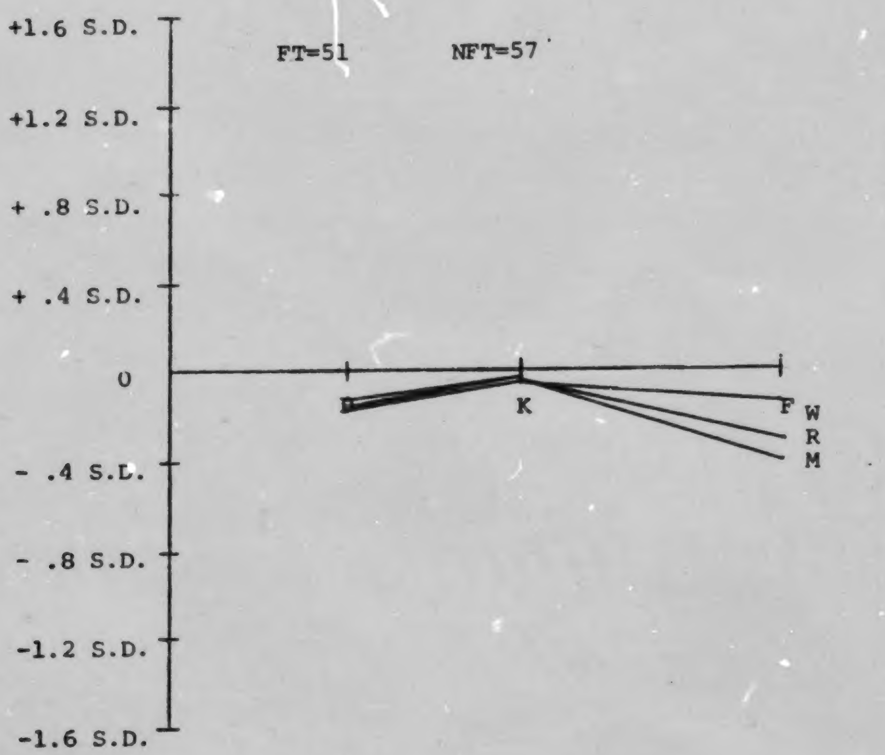
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	-.0859	-3.0090	-2.1511
S.E.	.0548	2.5797	3.0597
Two Year Difference (B)	-.0516	-2.5237	- .4316
S.E.	.0495	2.4503	2.7148
First Grade Difference (C)	-.0557	-2.6974	- .5898
S.E.	.0471	2.3709	2.6620
Outcome S.D.	.2030	9.4910	11.1967
P=(A-B)/S.D.	-.1690	- .0511	- .1536
K=(A-C)/S.D.	-.1488	- .0328	- .1394
F=A/S.D.	-.4232	- .3170	- .1921

Figure I-4 First Grade Outcome Profile for Far West Site NE

Note: All items marked by * are significant ($p < .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .0326	- 5.1682*	- 5.6223
S.E.	.0447	2.3292	3.3459
Two Year Difference (B)	.0080	- 3.5421	- 2.7660
S.E.	.0363	1.9878	2.6569
First Grade Difference (C)	- .0203	- 4.7114*	- 4.9170*
S.E.	.0314	1.8535	2.2617
Outcome S.D.	.2315	12.3040	17.5064
P=(A-B)/S.D.	- .1754	- .1322	- .1632
K=(A-C)/S.D.	- .0531	- .0371	- .0403
F=A/S.D.	- .1408	- .4200	- .3212

Figure I-5 First Grade Outcome Profile for Far West Site W2

Note: All items marked by * are significant ($p \leq .05$).

patterns under a broader range of conditions. At the same time, it is very important that a careful examination be made of the conditions which differentiate between successful and unsuccessful applications of the model. A model such as this one, which necessarily reflects the unique characteristics of each site in the construction of a "responsive environment", may not be expected to "work" under all conditions, and we should know why this is so. It is possible that a more robust version of the model can be developed or will emerge in the second or third grades and site-specific information can help a great deal in that process.

1.2 University of Arizona, Tucson Early Educational Model

The University of Arizona program focuses upon four general areas of development: language, reasoning abilities, motivation, and social arts and skills. The classroom environment is designed to reflect the child's home and community environment, so that skills and concepts may be learned in a natural, functional setting. One-to-one adult/child interactions and small group activities are utilized to individualize instruction and to help develop effective social interaction and communication skills.

Five sites met the criteria for inclusion in the evaluation of the University of Arizona model. These sites are scattered geographically and their students differ in a number of ways. Further, in four of the five sites the FT groups come from considerably poorer households than their NFT counterparts.

1.2.1 Site Descriptions

The five Arizona sites differ greatly in size and location. Two are located in the Northeast: a small town (NE1) and a large city (NE2). Two are located in the North Central region: a medium-sized city (NC1) and a large city (NC2). The remaining site is a small southern town.

In the small Southern community (S), the FT/NFT groups have different ethnic compositions (FT - 40% Black, 60% White; NFT = 98% White). On measures of parental education, family income, and father's occupation, the NFT group exceeds the FT group by a substantial margin. In addition, the NFT children in this community are far more apt to come from intact families (FT - 54%; NFT = 94%). (See Table VI-6, Vol. IIA.)

In the medium-sized North Central city (NC1), both FT and NFT children are predominantly White; in the large North Central city (NC2), both groups contain roughly the same proportion of Black and White children. In both communities, the FT children are much poorer than the NFT children. In NC1, both FT/NFT groups come from intact families; in NC2, the FT group has a smaller percentage of intact families (FT = 65%; NFT = 83%). (See Tables VI-7 and VI-8, Vol. IIA.)

In the two Northeastern communities both FT/NFT students are predominantly Black. In NE1 the FT group is well below the NFT group on two SES indices, mother's education and family income. However, in both groups more than half the children come from broken homes. (See Tables

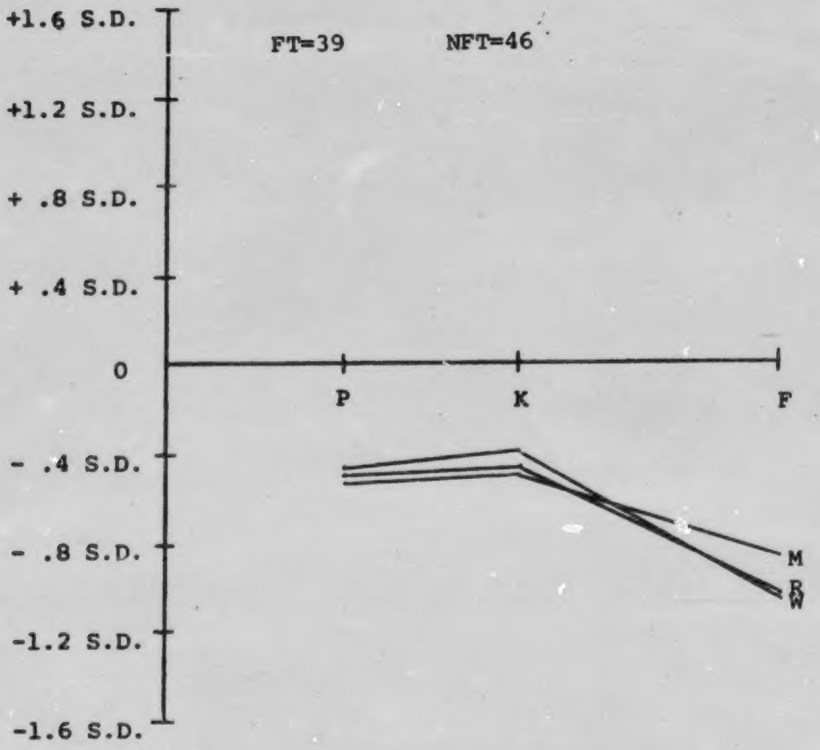
VI-9 and VI-10, Vol. II-A.) Only in the large Northeastern city, NE2, are the FT/NFT groups roughly equivalent on both SES and kindergarten entry tests.

1.2.2 Effects by Site

The Arizona sites show a consistent pattern on all outcome measures (Figures I-6 through I-10). During kindergarten the FT children in general maintain the level of their entry scores. In first grade, however, they fall substantially behind the NFT children. This pattern is observable in all sites on all outcome measures.

1.2.3 Summary

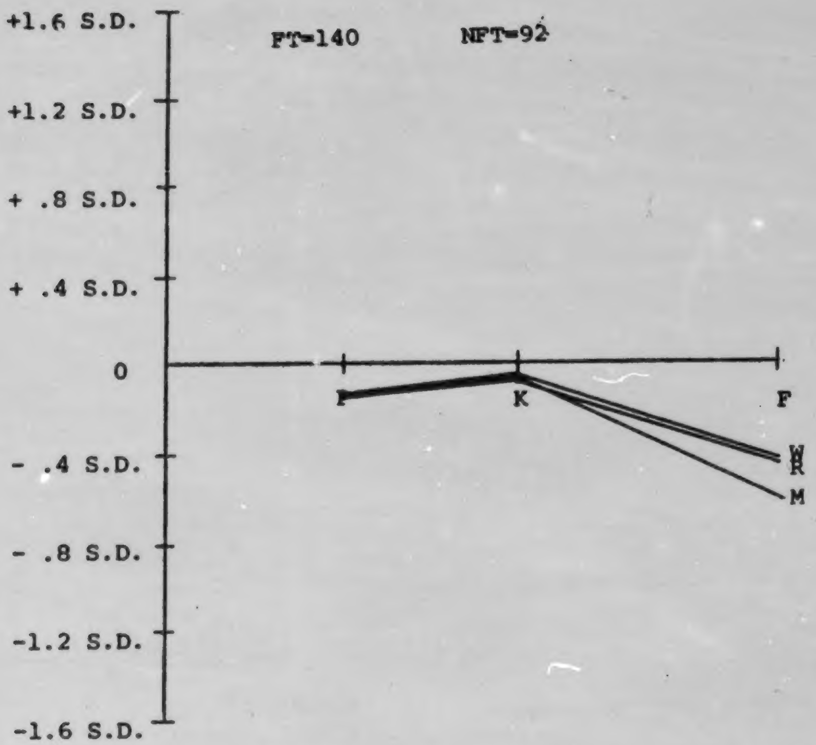
It is not clear how to account for these consistently negative effects. It is entirely possible that these effects are not uniform across all kinds of children; indeed, earlier analyses indicated that children with higher entry scores showed higher adjusted posttests than the NFT children. However, the clear pattern of negative longitudinal effects is a dominant finding which cannot easily be dismissed because of some isolated successes. It is obvious that the magnitude of negative findings is overstated because of the strong FT/NFT mismatch, the impact of which cannot be dealt with in the present analytic strategy. Much more important, however, is a set of anecdotal information which suggests that implementation of the Arizona curriculum has been beset by a great many problems. The most reasonable conclusion to be drawn from the present analysis is that a useful test of the effectiveness of this model has not yet been accomplished. Data from second and third grades are more likely to cast light on the effectiveness of the Arizona approach. The present negative findings are too far out of context to draw substantive conclusions at this point.



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .2918*	-11.8684*	-19.3333*
S.E.	.0485	2.6782	3.3138
Two Year Difference (B)	- .1672*	- 4.4306*	-10.1154*
S.E.	.0411	2.1129	2.6384
First Grade Difference (C)	- .1850*	- 4.9150*	-11.1424*
S.E.	.0336	2.0277	2.2624
Outcome S.D.	.2654	13.6007	17.9701
P=(A-B)/S.D.	- .4695	- .5469	- .5130
K=(A-C)/S.D.	- .4024	- .5112	- .4558
F=A/S.D.	- 1.0995	- .8726	- 1.0758

Figure I-6 First Grade Outcome Profile for Arizona Site S

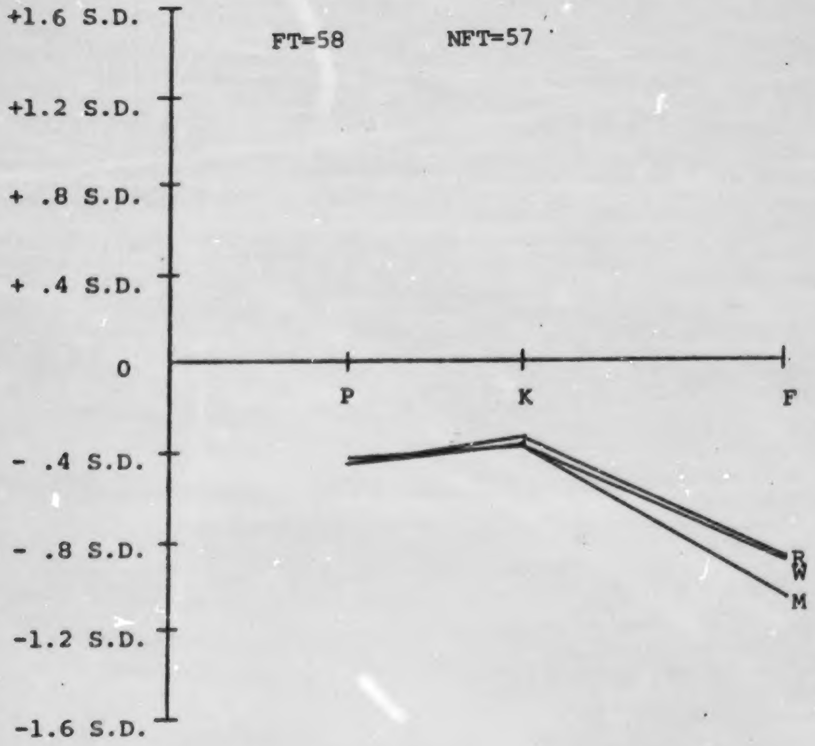
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .1006*	- 7.9062*	- 8.6043*
S.E.	.0310	1.5931	2.4361
Two Year Difference (B)	- .0696*	- 6.1329*	- 5.8025*
S.E.	.0237	1.1622	1.7464
First Grade Difference (C)	- .0888*	- 7.1100*	- 7.3880*
S.E.	.0220	1.0730	1.5856
Outcome S.D.	.2358	12.4623	18.5968
P=(A-B)/S.D.	- .1315	- .1423	- .1507
K=(A-C)/S.D.	- .0500	- .0639	- .0654
F=A/S.D.	- .4266	- .634[illegible]	- .4627

Figure I-7 First Grade Outcome Profile for Arizona Site NCl

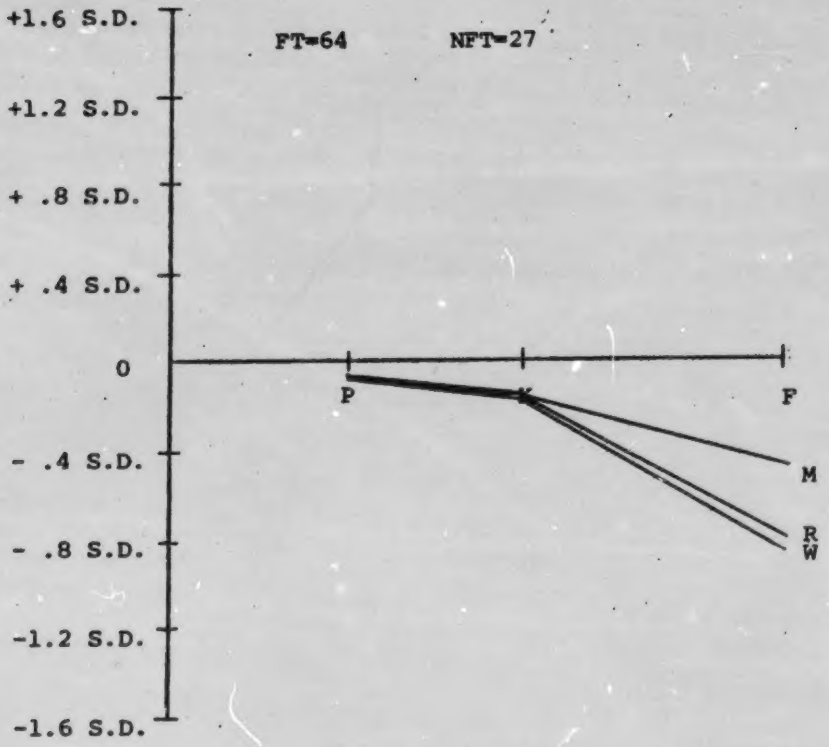
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .2244*	-14.7187*	-15.2680*
S.E.	.0435	2.1683	2.9078
Two Year Difference (B)	- .1069*	- 8.6466*	- 7.2705*
S.E.	.0355	1.8612	2.3992
First Grade Difference (C)	- .1347*	- 9.4223*	- 8.9235*
S.E.	.0321	1.8377	2.2363
Outcome S.D.	.2581	13.7331	17.3127
P=(A-B)/S.D.	- .4552	- .4422	- .4619
K=(A-C)/S.D.	- .3475	- .3857	- .3665
F=A/S.D.	- .8694	- 1.0718	- .8819

Figure I-3 First Grade Outcome Profile for Arizona Site NC2

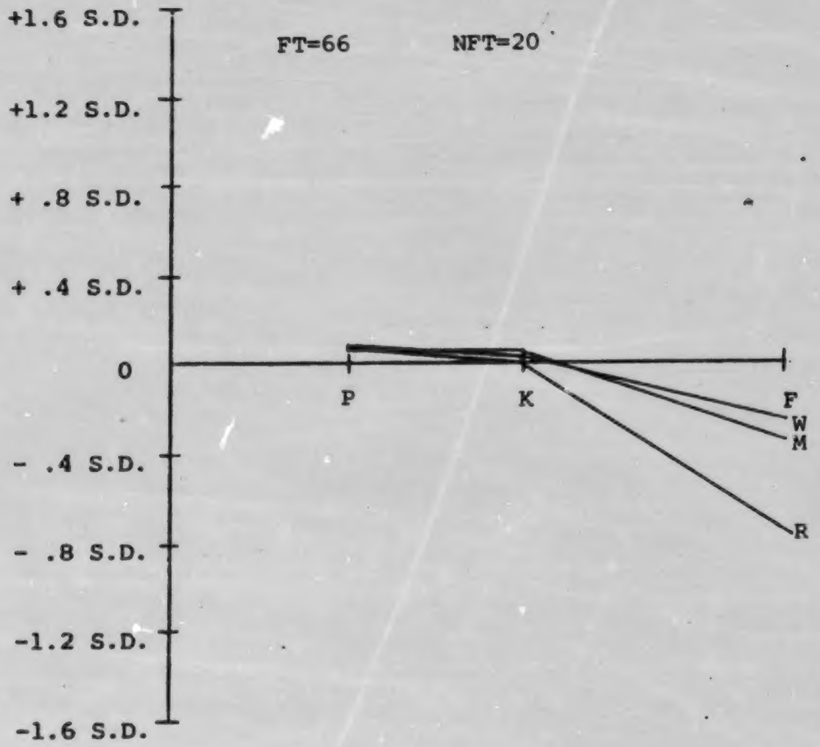
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .2320*	- 5.2714*	-12.9219*
S.E.	.0563	2.5436	3.3211
Two Year Difference (B)	- .2085*	- 4.3099*	-11.6337*
S.E.	.0449	2.0707	2.6603
First Grade Difference (C)	- .1812*	- 3.3440	-10.3018*
S.E.	.0427	2.0319	2.5936
Outcome S.D.	.2663	11.2852	15.5675
P=(A-B)/S.D.	- .0882	- .0852	- .0827
K=(A-C)/S.D.	- .1908	- .1708	- .1683
F=A/S.D.	- .8712	- .4671	- .8300

Figure I-9 First Grade Outcome Profile for Arizona Site NE1

Note: All items marked by * are significant ($p < .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .0619	- 3.1955	-11.6470*
S.E.	.0630	2.2569	3.5931
Two Year Difference (B)	- .0811	- 3.8600*	-12.7050*
S.E.	.0491	1.8017	2.8714
First Grade Difference (C)	- .0678	- 3.6642*	-11.7928*
S.E.	.0474	1.8093	2.7232
Outcome S.D.	.2469	8.8939	14.8433
P=(A-B)/S.D.	.0778	.0747	.0713
K=(A-C)/S.D.	.0239	.0527	.0098
F=A/S.D.	- .2507	- .3593	- .7847

Figure I-10 First Grade Outcome Profile for Arizona Site NE2

Note: All items marked by * are significant ($p \leq .05$).

1.3 Bank Street College Approach

The Bank Street approach is designed to change the school system to meet the developmental needs of children. Heavy emphasis is placed on teacher training to help teachers organize the classroom environment for self-directed learning and to plan events to meet the needs of children. The individualized, flexible curriculum is designed not only to help children acquire basic skills, but also to help children master how to learn. Creativity and self-expression are also important program goals.

An evaluation of the Bank Street program is severely handicapped by drastic losses of sites and children. Of the four sites available for study in the first grade year, two, both big cities, had inordinately high rates of attrition. The numbers of children for whom complete data are available at these two sites are not sufficient to justify analysis. The two remaining sites with large enough numbers to permit site level analysis are both New England cities with demographic and educational characteristics unique among the Bank Street sample of sites. The effects found for this limited and unique subset of sites cannot be taken as a full assessment of the impact of this model.

1.3.1 Site Descriptions

The two New England sites, one (NE1) a small city of 15,000 and the other (NE2) a middle sized city of 100,000 are almost exclusively white. The smaller city is largely Yankee in origin whereas NE2 has a large non-English speaking minority group.

An important characteristic of NE1 is the very sharp SES differences between the FT and the NFT groups. The magnitude of the income differential between these two groups is about the largest in the national sample. This is in part the result of the fact that the NFT group has one of the highest family incomes in the national sample. The FT group has substantially fewer intact families and is significantly behind the NFT group on both the WRAT and the PPVT at the beginning of kindergarten. In addition, there is anecdotal evidence that candidates for Follow Through are selected from among the lowest achieving children in the kindergarten classes. The very sharp social stratification which exists in this small town appears to be reflected in the FT/NFT differences, which makes the task of Bank Street and of the evaluation very difficult.

The second site is very different from the first. Here, both the FT and the NFT groups are from the low end of the income distribution and like NE1 are distinguished primarily by the number of intact families (FT = 57%, NFT = 81%). Both have substantial numbers of non-English speaking families although the FT group exceeds the NFT group on the PPVT.

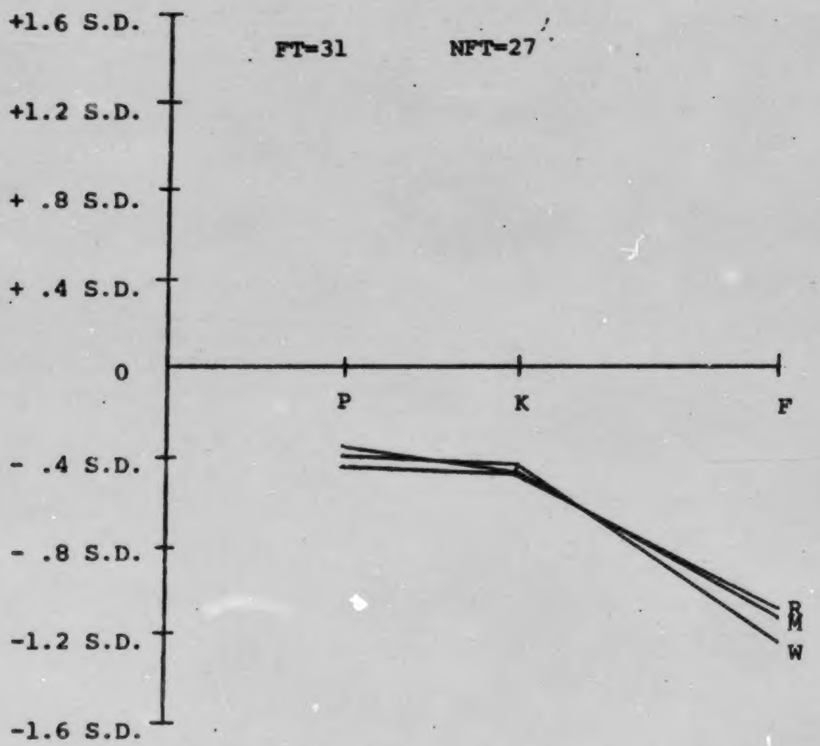
1.3.2 Effects by Site

Two different patterns are appparent in these sites. The smaller city (NE1) shows a maintenance of the relative FT/NFT growth during the kindergarten year and then a much slower growth among the FT than the NFT children during the first grade. This is the case on all outcome measures.

In the larger city (NE2), relative losses for the FT children are apparent during the kindergarten year, and a leveling off of the gap in Word Analysis and Reading during the first grade. The shift in the relative growth during first grade is even more apparent in math where the FT group has so increased the relative rate of achievement that the gap between them and the FT children has disappeared entirely.

1.3.3 Summary

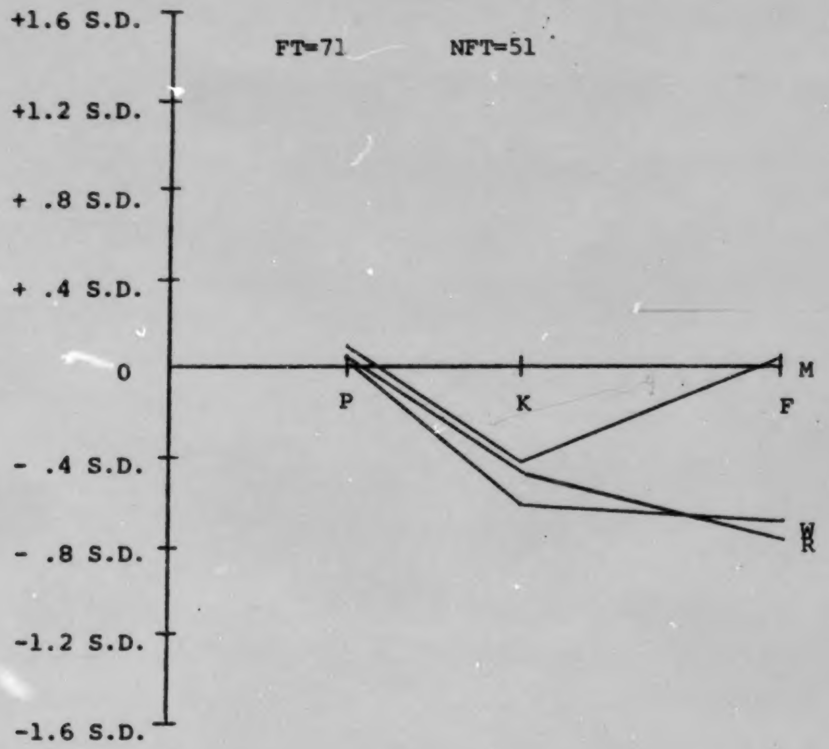
The very large attrition observed for this Sponsor, and the large degree of mismatch between the FT and NFT groups makes interpretation of these findings extremely tentative. The evidence up to this point in the evaluation suggests that the gap between the very disparate groups in NE1 has not been reduced. Deficit appears to be accumulating although this might well be a function of the extreme mismatch (and the tendency of the attrition to maintain or increase the mismatch). Where the mismatch in performance is low (albeit the mismatch in social status is likely to be high), the initial loss appears to be arrested in the first grade. The Bank Street model may require several years of implementation before effects as measured by achievement tests can emerge. At least this may be true in those instances in which very large initial differences are not found. Future data will clarify this issue.



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .3750*	-14.0872*	-22.2389*
S.E.	.0619	2.6479	4.3104
Two Year Difference (B)	- .2552*	- 9.7270*	-13.1739*
S.E.	.0484	2.3523	3.2274
First Grade Difference (C)	- .2447*	- 8.6072*	-13.1168*
S.E.	.0511	2.4409	3.4210
Outcome S.D.	.2998	12.2330	19.7137
P=(A-B)/S.D.	- .3996	- .3564	- .4598
K=(A-C)/S.D.	- .4346	- .4480	- .4627
F=A/S.D.	- 1.2508	- 1.1516	- 1.1281

Figure I-11 First Grade Outcome Profile for Bank Street Site NE1

Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .1707*	.0295	-12.0696*
S.E.	.0424	1.9681	2.5873
Two Year Difference (B)	- .1775*	- .9928	-12.6170*
S.E.	.0362	1.6616	2.1728
First Grade Difference (C)	- .0170	4.5748*	- 4.8879*
S.E.	.0383	1.9063	2.4605
Outcome S.D.	.2453	10.6779	15.2572
P=(A-B)/S.D.	.0277	.0957	.0359
K=(A-C)/S.D.	- .6266	- .4257	- .4707
F=A/S.D.	- .6959	.0028	- .7911

Figure I-12; First Grade Outcome Profile for Bank Street Site NE2

Note: All items marked by * are significant ($p \leq .05$).

1.4 University of Oregon

The University of Oregon approach is designed to teach children the basic skills of reading, arithmetic, and language. It is based on the assumption that disadvantaged children can perform at "normal" levels of achievement when the instructional program builds on the skills children bring with them to school. The curriculum materials are programmed and sequenced. Teaching techniques with heavy emphasis on structured small group instruction, quick paced question-and-answer periods, and the use of positive reinforcement to shape behavior are stressed.

Three sites are included in this analysis of the University of Oregon model. They are all cities of approximately 200,000 people, somewhat diverse geographically, although they do not include any of the Southern communities with which this Sponsor works. In addition, the FT and NFT children in our analytic sample differ in their demographic and kindergarten entry level characteristics and these differences are unique to some of the sites. It should be kept carefully in mind when we consider the results of the analyses that it is very difficult to separate site and demographic effects from model effects.

1.4.1 Site Descriptions

The sites available for analysis fall within two distinct categories. The first category includes two North Central sites (NC1, NC2) in which the vast majority of both FT and NFT children come from low income Black families. (See Tables VI-13 and VI-14, Vol. II-A.) In both communities the FT families have substantially lower incomes than the NFT families which may be a reflection of the fact that there is a significantly lower percent of intact FT families and therefore fewer potential wage earners. In any case, associated with a much higher percentage of single parent homes, the FT children come from a considerably more precarious economic position than the NFT children.

The relatively lower status of the FT than the NFT families in these sites is further reflected in the much higher percents of FT children drawn into Head Start and other preschool programs. Thus, in NC1, 93% of the FT children attended preschool, whereas only 63% of the NFT children enrolled in such programs. Site NC2 showed FT/NFT percentages of 94 and 20 respectively in preschool prior to entrance into kindergarten. Outreach to the more economically disadvantaged is likely to be operating in these communities, and the groups which ultimately reached the Follow Through classes appear to have been the target of the Outreach.

One factor which distinguished the two sites within this category is the pattern of kindergarten entry scores in the WRAT. NC1 has an overall mean of approximately 25 whereas NC2 has an overall approximate mean of 30. However, in both cases, the more economically disadvantaged FT children significantly exceed the NFT children! (NC1: FT = 27.96, NFT = 23.73; NC2: FT = 33.69, NFT = 27.57). At this point it is not clear if this is an example of a preschool effect, or a self selection effect. That is, the much higher rates of enrollment in preschool among the FT groups may reflect as much an intention on the part of FT mothers toward upward mobility in deliberately enrolling their children in preschool as it may reflect the outreach efforts of the preschool staff. In any case, the FT children, as they entered kindergarten, appear to have established a momentum of achievement which is not often seen in children from this economic category. This momentum is not found in all cognitive domains since there is no difference among any of the groups on the entry scores of the PPVT. Nevertheless, in these two communities (one of which appears to be a generally higher achieving locale than the other), the FT children, despite their more diminished economic status, appear to be entering public schools with better achievement skill than the NFT.

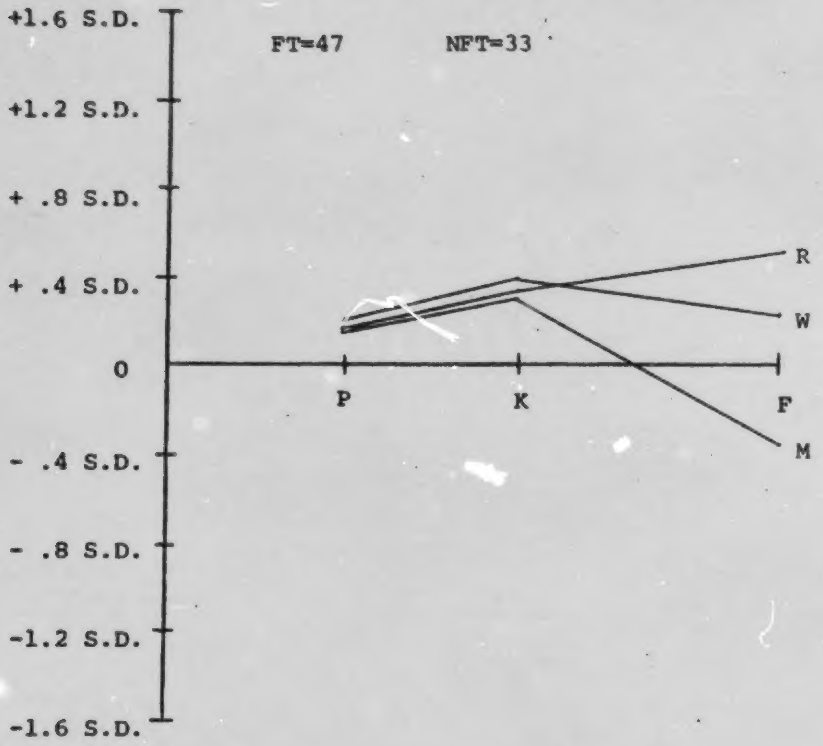
It is important to note that the NFT children are also clearly economically disadvantaged. Their schools all receive more Title I allocations per pupil than do the FT schools (which may reflect the Follow Through support), and in the case of NC2, the number of man days of professional support services per NFT child is substantially greater than that for FT children. Clearly the NFT children appear to the school officials of their communities to be in need of support although their parents, or the preschool staff, did not feel they were in need of preschool experiences. Obviously the data available do not allow for a precise estimate of the extent of the match/mismatch between these groups.

The second category includes just one large northeastern city. (See Table VI-15, Vol. II-A.) Here 76% of the FT children are Black, in contrast to 58% of the NFT children. Further, as measured by all available indicators, the FT children come from families with lower SES. Income, parental education and percent of intact families are all significantly lower in the Ft group. Further, every FT child attended a preschool prior to kindergarten whereas only 15% of the NFT children had such experiences. In addition, FT children attend public schools which received substantial amounts of Title I funds in addition to Follow Through support, whereas the NFT schools in this site receive virtually no Title I support. There appears to be a sharp SES as well as ethnic membership difference between these two groups not found in the sites of the first category.

Upon entry into kindergarten, administration of the WRAT revealed significant differences between the FT and NFT groups in favor of the latter, which further distinguished this category from the previous one. If preschool can account for the differences in performance in the children in the first category of sites, it does not seem to be able to in this category. This is true despite the fact that most children in this site (NE) are generally higher achieving than those in either of the first two sites. It is possible, of course, that the rather high achieving FT group in NE (29.97), a mean which is higher than 3 of the 4 groups in the first category), received these scores with the help of their preschool experiences. If this was the case, the original (pre-preschool) achievement levels must have been astonishingly lower than the NFTs' original levels, which implies a severe mismatch. In any case, the NE site is composed of FT children who are well behind their NFT comparisons on the WRAT, as well as on all SES predictors of school performance. This site appears to be similar to other large urban centers in which there are large concentrations of low achieving, low income children. Unfortunately, the comparison group appears to be higher achieving and higher income which makes the comparison very difficult.

1.4.2 Effects by Site

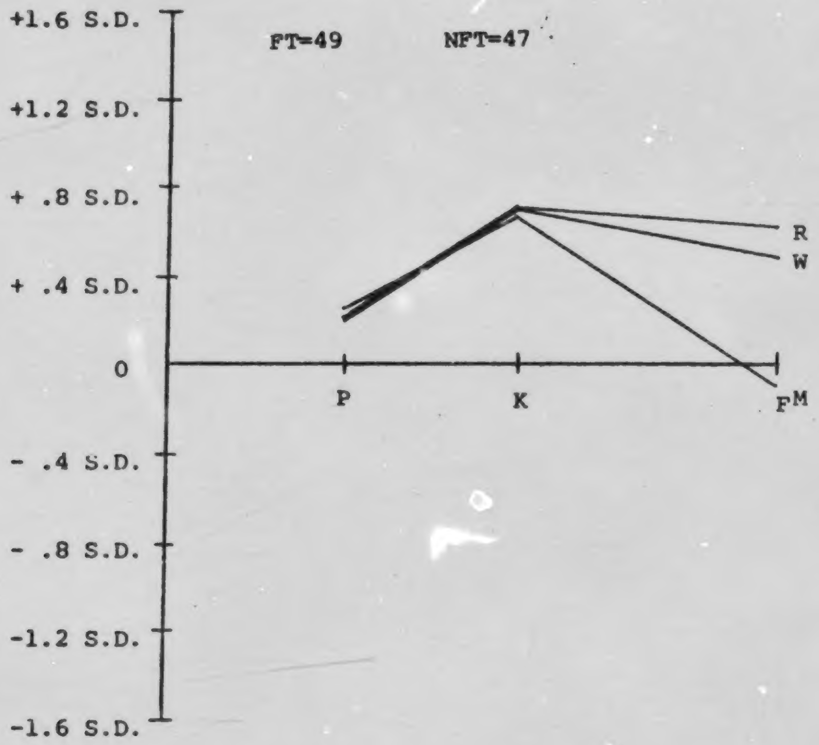
Figures I-13, 14, 15 summarize the findings for Sites NC1, NC2, and NE respectively.



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0495	- 3.9400	8.1103*
S.E.	.0481	2.4239	3.4544
Two Year Difference (B)	.0080	- 5.4614*	5.6480
S.E.	.0433	2.2306	2.9360
First Grade Difference (C)	- .0345	- 7.2317*	2.7648
S.E.	.0412	2.1961	2.7933
Outcome S.D.	.2120	10.7829	15.6387
P=(A-B)/S.D.	.1958	.1411	.1574
K=(A-C)/S.D.	.3962	.3053	.3418
F=A/S.D.	.2335	- .3654	.5186

Figure I-13 . First Grade Outcome Profile for Oregon Site NC1

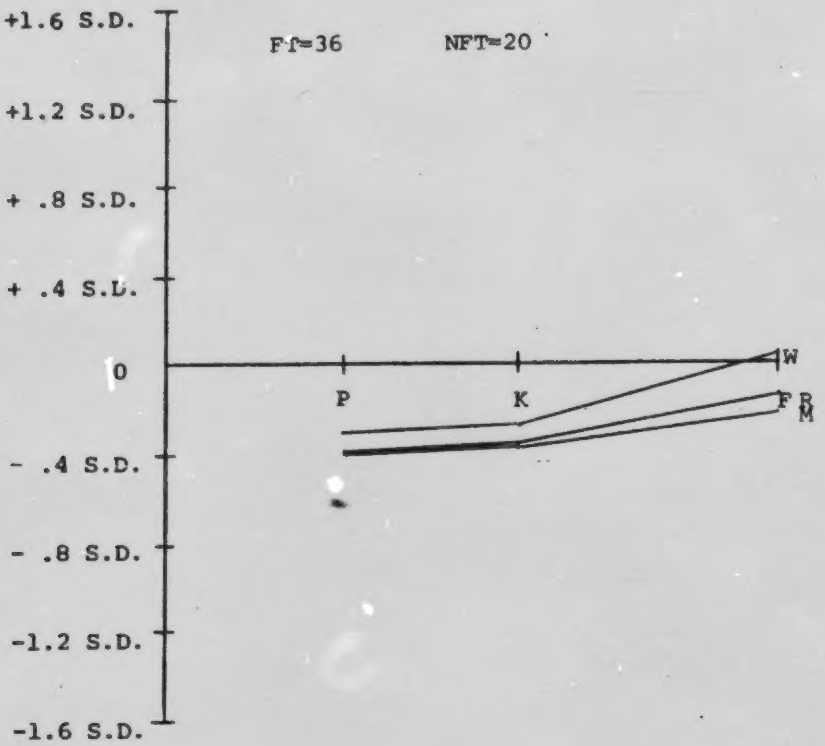
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0963*	- 1.0261	10.0426*
S.E.	.0399	2.3501	3.0945
Two Year Difference (B)	.0560	- 3.9120	6.7751*
S.E.	.0397	2.2614	3.0132
First Grade Difference (C)	- .0464	- 8.6785*	- 1.3409
S.E.	.0429	2.5305	3.2231
Outcome S.D.	.2006	11.4615	15.8988
P=(A-B)/S.D.	.2009	.2518	.2055
K=(A-C)/S.D.	.7114	.6677	.7160
F=A/S.D.	.4800	- .0895	.6316

Figure I-14 First Grade Outcome Profile for Oregon Site NC2

Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0077	- 2.8000	- 2.3667
S.E.	.0637	3.5161	4.8238
Two Year Difference (B)	.0780	2.3979	4.5271
S.E.	.0614	3.1034	3.9067
First Grade Difference (C)	.0695	1.9648	4.0384
S.E.	.0525	2.6462	3.4610
Outcome S.D.	.2262	12.5655	17.1770
P=(A-B)/S.D.	- .3108	- .4137	- .4013
K=(A-C)/S.D.	- .2732	- .3792	- .3729
F=A/S.D.	.0340	- .2228	- .1378

Figure I-15 First Grade Outcome Profile for Oregon Site NE

Note: All items marked by * are significant (p ≤ .05).

The pattern of effects can be roughly distinguished according to the two categories of sites identified above. The first category is characterized by rather impressive gains during the kindergarten year, and significantly slower gains than the NFT group during the first grade. Math achievement shows the pattern with particular clarity. In both sites the FT children decreased their rate of achievement in Math relative to the NFT rate very significantly in first grade, after a kindergarten year in which they achieved at a much more rapid rate than the NFTers. Much the same is true for the Word Analysis subtest, and it is true to a lesser extent in reading in the NC2 site. Reading in the NC1 site continued to improve in the first grade after a relatively greater growth rate in kindergarten. This latter finding is a true cumulative effect: the gap between FT and NFT groups in reading increased by the end of kindergarten, and still further increased by the end of first grade. The contrast of this finding to all other findings for the sites in this category is striking: in all other cases, the FT/NFT gap diminished (and in some cases reversed so that the NFT was ahead) by the end of first grade after a kindergarten year in which the FT favoring gap increased.

The pattern of effects for the NE site in the second category is very different from the one just described. Here, the FT children started out well behind the NFT children on all three measures, diminished the gap very slightly at the end of kindergarten and continued to diminish the gap very slightly by the end of first grade. Thus the differences between FT and NFT at the end of first grade is about zero on all measures whereas it was significantly large at the beginning of kindergarten. Most of the relative gain of the FT children was accomplished during the first grade rather than during kindergarten which is just the opposite pattern to that observed in the sites of category one.

1.4.3 Summary

The effects which have been produced by the University of Oregon to date (as observed in the limited subset of their sites which are available for analysis) might best be described according to two dimensions: the time of impact and the relative status of the FT/NFT groups. When dealing with (Black) children who are relatively higher achieving than their (Black) comparisons, the kindergarten effects are very strong and are followed by almost as strong first grade losses. When dealing with (Black) children who are well behind their (White) comparisons, the kindergarten and first grade effects are small and cumulative so that the gap appears to be closed by the end of first grade. These FT children are now at the same level, relative to their NFT comparisons, that the children in category one (i.e., sites NC1 and NC2) achieved by their entrance into kindergarten. It remains to be seen if the future performance of the NE children begins to look like the pattern produced by the children in NC1 and NC2, or if the NE children continue their cumulative gain.

1.5 University of Kansas Behavior Analysis Approach

The primary objective of the University of Kansas approach is to facilitate the child's mastery of basic skills particularly in reading and arithmetic through the establishment of a "token economy" within the classroom. Based on basic principles of behavior modification, the token exchange system is designed to provide systematic, positive reinforcement for desired behavior. The tokens, which are given as rewards for successful completion of tasks, may later be exchanged for desired activities. Within the "token economy" environment, programmed instructional materials are used to teach basic skills.

This analysis of the Kansas program examines data from two types of sites. The first category includes three large cities in the Northeast, North Central, and Southern regions of the country in which the children attending the FT and NFT schools are almost exclusively lower SES Black children. In the second category there is one rural site in which the children are predominantly White and higher in SES and kindergarten entry level than the children in the urban sites. The effectiveness of the program may depend on the site (urban or rural), on the ethnic background of the children, on neither of these, or on both. In these data, however, the effects, if any, of site and ethnic background are confounded and no conclusions about their separate effects can be drawn. This limitation must be borne in mind throughout the following discussion.

1.5.1 Site Descriptions

1.5.1.1 Urban Sites

These three sites are located in the Northeast (NE), North Central (NC2) and Southern (S) region of the country. The demographic properties of the FT and NFT groups show an almost identical pattern in all three sites as seen in Tables VI-16 through VI-18 (Vol. II-A). The ethnic composition in all three is at least 95% Black in both FT and NFT groups. Preschool attendance is high in all groups in all sites (mean FT = 76%; mean NFT - 86%), and Title I expenditures are high as well. The FT/NFT differences are almost identical in all sites: FT groups are somewhat lower in total family income, and there is a smaller percentage of intact families in the FT groups in each site. It is likely that father absence accounts for the differences in income since there are fewer wage earners in these families. This suggests that the social status differences between the FT and NFT groups are not as great as the income differential implies, although the difference in family structure does represent a real psycho-educational difference between the groups.

1.5.1.2 Rural Site

We turn now to the second category of Kansas sites which includes NC1, a small midwestern town. Here the sample has predominantly intact families (FT = 85%; NFT = 76%), and fewer children attend preschool (FT = 66%; NFT = 55%), compared to the urban sites. (See Table VI-19, Vol. II-A). Despite these demographic similarities within this site, the FT children rank consistently higher on all SES indicators than the NFT group. FT family income, in particular, is substantially higher. These status differences are also reflected in sharp differences between FT and NFT on initial entry level. The FT group scores almost a full standard deviation higher than the NFT on the Fall kindergarten WRAT.

It seems likely that this site accounts for one of the issues raised in the earlier studies involving the Kansas model. In the analysis of the kindergarten data for this Sponsor reported in Volume 1A of "Education as Experimentation," pretest differences in favor of the FT group were identified. Such differences pose severe problems for analysis of covariance and consequently threatened the interpretability of the analyses of the Kansas data. Two hypotheses were generated to account for the differences. The first was that the Sponsor had achieved a very early impact of the model and that the differences were in fact treatment effects. The second hypothesis was that the two groups were truly different upon entry into the program and that unknown selection procedures, which confound treatment effects, were responsible for the assignment of the higher achieving (and higher SES) children to FT and the lower achieving children to NFT in the communities in which this sponsor is working. Although the data available at that time tended to support the latter hypothesis, and tended to reject the hypothesis of early treatment effects, the issue was not resolved.

The present data cast some new light on this issue. The overall analysis of the Kansas data, in which all children from all sites are aggregated to the Sponsor level, reveals the same FT/NFT differences on kindergarten pretest found in the comparable analysis run for the earlier report. However, these data can be disaggregated to the site level, and here it is possible to see that the source of the difference is to be found primarily in site NC1. Thus the overall WRAT entry means for FT and NFT respectively are 27 and 24, representing a significant difference in favor of the FT group. However, site NC1 shows the only significant difference on mean entry level of the four sites included in the analysis (FT = 33; NFT = 24) and, when the children in this site are removed from the overall mean, the difference between the remaining FT and NFT means is reduced to 25 and 24, respectively, which is not significant. It is likely, therefore, that the SES differences of the children in this site are the source of the FT/NFT differences in the overall analysis.

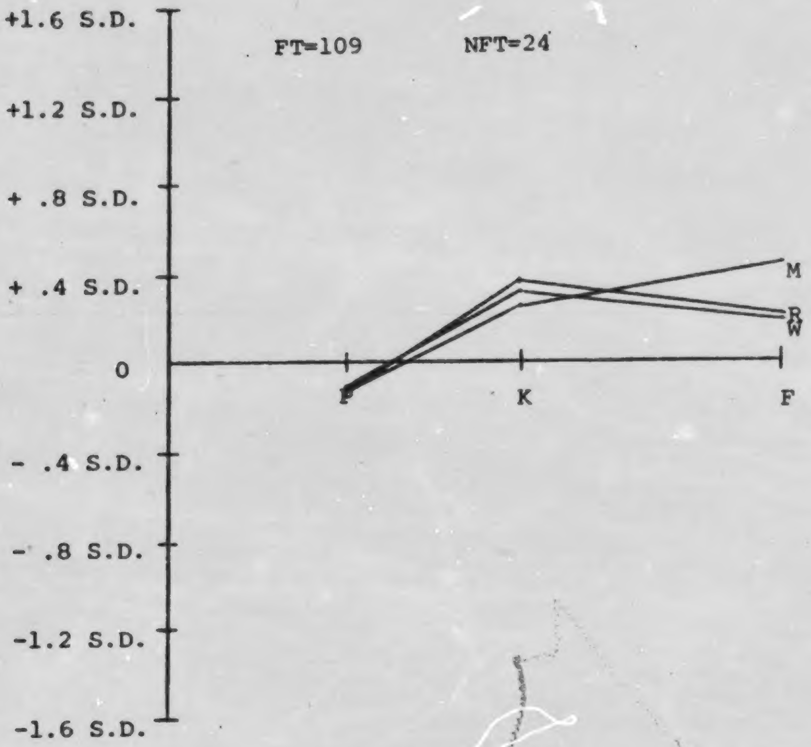
1.5.2 Effects by Site

1.5.2.1 Urban Sites

The reading results for each of the sites are similar: strong gains in kindergarten and a sharp reduction in the gains relative to NFT in the first grade (Figures I-16 through I-18).

Examination of the effects of the program on Word Analysis skills show a very variable pattern. Although the FT children at all three sites showed much higher adjusted achievement levels in this area at the end of kindergarten, the children at only one site continue to acquire these skills at the same relative rate by the end of first grade. Children at another site are able to maintain the same position at the end of first grade that they achieved at the end of kindergarten, but they are not able to widen the gap as they did during their first year in the program. The children in the third site show some losses relative to the NFT group although they, too, are achieving at a higher level than their comparison group at the end of the year.

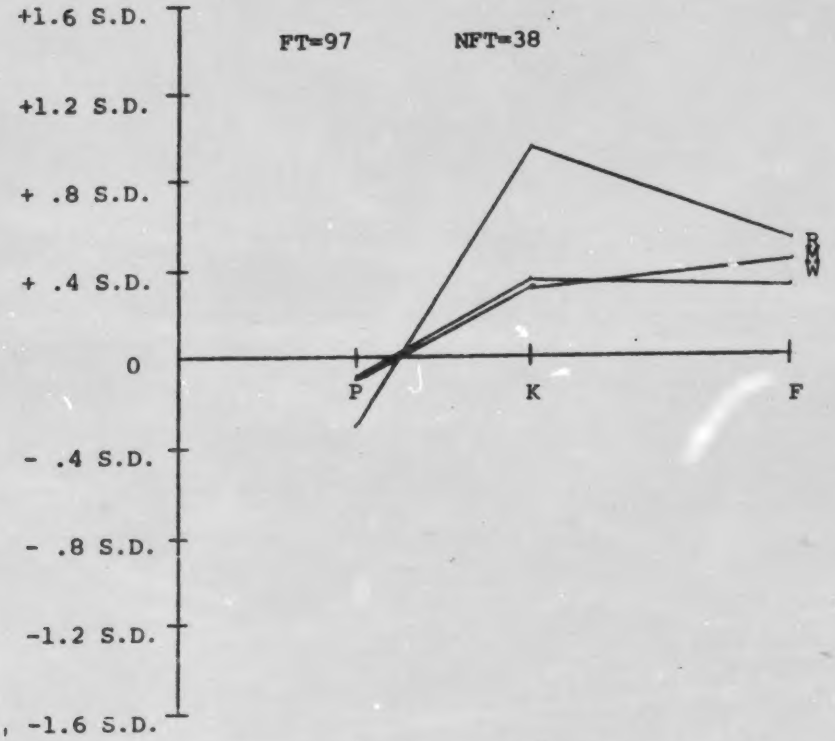
The results in Mathematics are quite different from those in Reading and Word Analysis and indicate consistent gains in first grade over and above those produced in kindergarten. There is a clear picture of



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0387	4.6005*	2.8567
S.E.	.0514	2.3018	3.4075
Two Year Difference (B)	.0622	5.6869*	4.5226
S.E.	.0442	1.9576	2.8790
First Grade Difference (C)	- .0330	2.0091	- 2.7343
S.E.	.0407	1.8795	2.4773
Outcome S.D.	.2278	10.3237	15.0954
P=(A-B)/S.D.	- .1032	- .1052	- .1104
K=(A-C)/S.D.	.3147	.2510	.3704
F=A/S.D.	.1699	.4456	.1892

Figure I-16 First Grade Outcome Profile for Kansas Site NE

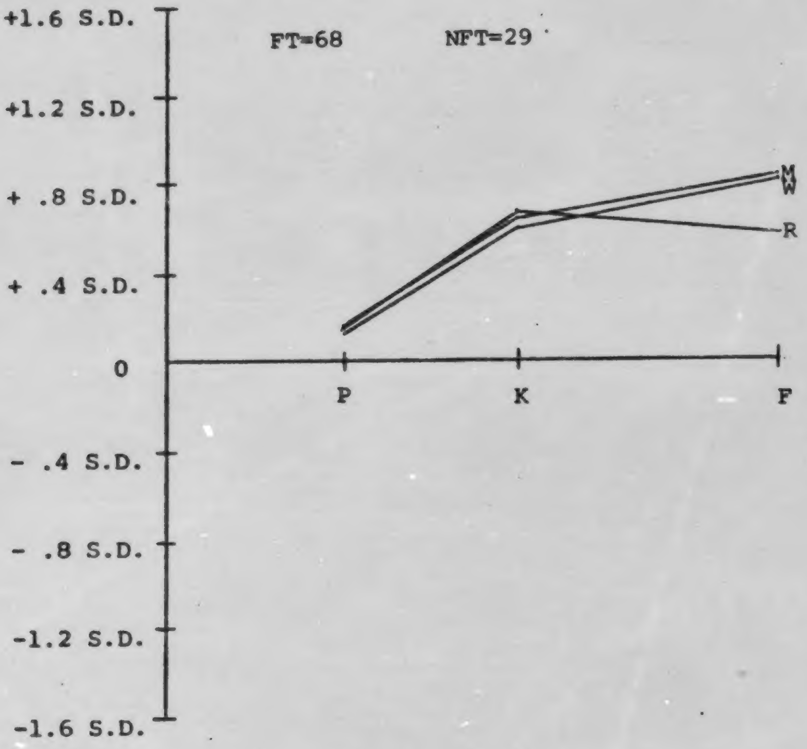
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0772	5.0038*	3.2282
S.E.	.0453	2.1332	3.0627
Two Year Difference (B)	.0997*	6.1824*	5.0768*
S.E.	.0351	1.7945	2.4463
First Grade Difference (C)	- .0073	1.4199	- 2.4730
S.E.	.0292	1.6148	2.0156
Outcome S.D.	.2384	11.3325	6.0102
P=(A-B)/S.D.	- .0944	- .1040	- .3076
K=(A-C)/S.D.	.3544	.3162	.9486
F=A/S.D.	.3238	.4415	.5371

Figure I-17 First Grade Outcome Profile for Kansas Site NC2

Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.2311*	[illegible].7464*	9.7759*
S.E.	.0572	2.8496	3.5972
Two Year Difference (B)	.1990*	9.4978*	7.2129*
S.E.	.0485	2.2935	2.8997
First Grade Difference (C)	.0634	2.8739	- 1.4218
S.E.	.0463	2.1619	2.6932
Outcome S.D.	.2778	13.8774	16.7499
P=(A-B)/S.D.	.1156	.1620	.1530
K=(A-C)/S.D.	.6037	.6393	.6685
F=A/S.D.	.8319	.8464	.5836

Figure I-18 First Grade Outcome Profile for Kansas Site S

Note: All items marked by * are significant ($p \leq .05$).

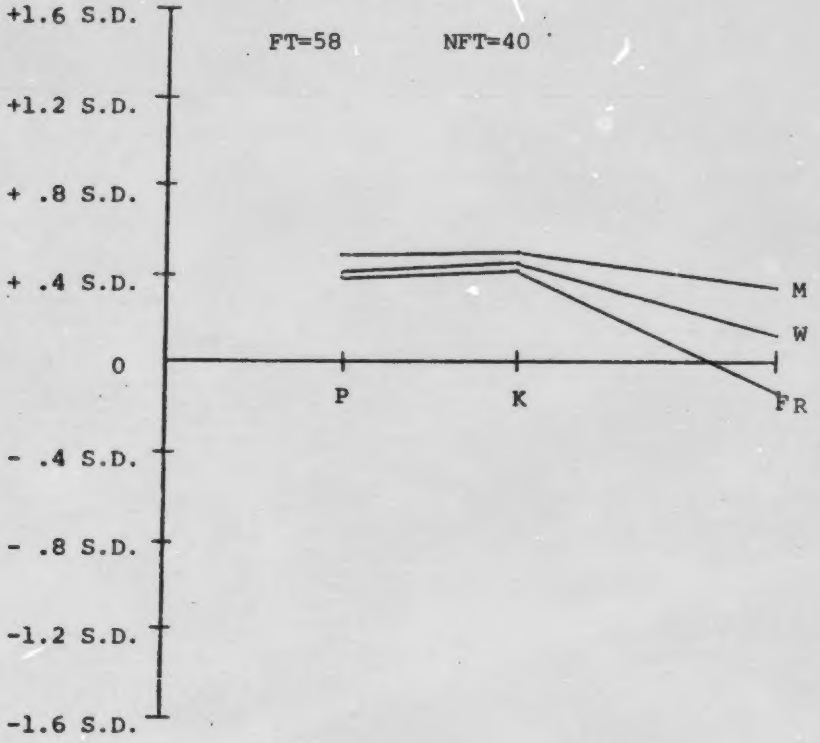
accumulation of effects over the grades and an indication that the Kansas program is able to help substantially the acquisition of math skills in children in these sites.

1.5.2.2 Rural Site, Effects of Program Within Site

The most impressive findings for this site, in comparison with the urban sites, are losses during first grade in all of the academic areas. The drop in reading skills relative to the NFT group is so great that the FT group is achieving at a lower level than the NFT at the end of the first grade when they were at a substantially higher level at the end of kindergarten. The second impressive finding at this site is the loss relative to NFT in math when the urban sites were all showing a gain. The Word Analysis effects are also much smaller in the FT than in the NFT group, although the gap has not been fully closed on this measure.

1.5.3 Summary

To summarize, it is clear that despite the very great variability in types of sites, reading effects are down at all of them. This suggests that the Kansas program may be having difficulty in dealing with the developmental tasks in reading faced by most first graders. First graders are moving away from the easily reinforced word and letter identification tasks and are moving toward the less easily reinforced search for meanings and comprehension. There is some evidence in the math results which also indicates that differences in the nature of the responses required in various tasks contribute to the kinds of gains shown by the children working in this program. There is reason to believe that at all the Kansas sites the largest source of math gains are in math computation and not in math problem solving. (This is true even at the rural site where the total math gains are much smaller in the first grade than they were in kindergarten.) These specific computational skills are more easily reinforceable than the steps involved in problem solving, and this difference may be reflected in the efficacy of the reinforcement model at the first grade. Such considerations may also



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0327	4.3836	- 2.4526
S.E.	.0508	2.6080	3.6151
Two Year Difference (B)	- .0718	- 1.8442	- 9.4364*
S.E.	.0416	1.8900	3.0640
First Grade Difference (C)	- .0777*	- 2.1274	- 9.8801*
S.E.	.0360	1.6008	2.6298
Outcome S.D.	.2464	12.8082	17.5404
P=(A-B)/S.D.	.4241	.4862	.3982
K=(A-C)/S.D.	.4480	.5083	.4234
F=A/S.D.	.1327	.3422	- .1398

Figure I-19 First Grade Outcome Profile for Kansas Site NC1

Note: All items marked by * are significant ($p \le .05$).

explain the diminished gains in the reading test, but they do not account for the variable findings in the word analysis measure. Clearly, the nature of the impact of the Kansas model is still not fully understood. The most that can be said at this point in the study is that the Kansas program is more effective among the urban sites with lower achieving child children than in the rural site with higher achieving children.

1.6 High Scope Foundation, Cognitively Oriented Curriculum

The High Scope classroom environment may be described as "open," with an emphasis on active exploration, manipulation, and discovery. Within this open framework, the instructional approach is systematic and planned. The cognitively oriented curriculum is highly Piagetian. The ultimate goal is to develop in children the thinking skills they will need throughout their school years and adult lives.

Included in this analysis of High Scope effects are six geographically diverse sites with diverse characteristics. In four of the sites, the FT/NFT groups are fairly well matched on SES and entering achievement; in two, they are severely mismatched.

1.6.1 Site Descriptions

1.6.1.1 Closely Matched Sites

The High Scope FT and NFT comparison groups are fairly closely matched in four of the six sites in this analysis. These sites are geographically scattered; two are large cities, one in the Northeast (NE) and the other in the North Central region (NC). The other two sites are in the West, one a middle-sized city (W1), the other a rural town (W3). In these sites, the FT and NFT groups are very similar on family structure characteristics and on all SES indicators except income level, which is somewhat lower in the FT groups. The FT and NFT children at these sites are predominantly Black and/or Spanish-surnamed children. The groups show no significant difference on kindergarten entry tests. (See Tables VI-20 through VI-23, Vol. II-A.)

1.6.1.2 Mismatched Sites

Two of High Scope's sites show a severe mismatch between FT and NFT. One of these sites is a small Southern city (S) and the other is a middle-sized Northwestern city (W3). In site S, the FT and NFT groups show no significant difference on parents' level of education but the groups are different on all other indicators. The income level in FT families is lower than in NFT families and the FT group has considerably fewer intact families.

The FT group is nearly all Black while the majority of NFT children are White. Finally, the groups show no difference on the kindergarten entry WRAT but a substantial difference in favor of the NFT group on entry PPVT. (See Table VI-24, Vol. II-A.)

Site W3 (Table VI-25, Vol. II-A) displays a similar disadvantage in the FT group. However, the mismatch appears to be more severe. The FT group is predominantly Black, while the NFT group is a mixture of White, Oriental and Black children. The FT group is lower than NFT on all SES indicators except the level of mother's education, and has considerably fewer intact families. The FT groups also show a substantial disadvantage on both the kindergarten entry WRAT and PPVT.

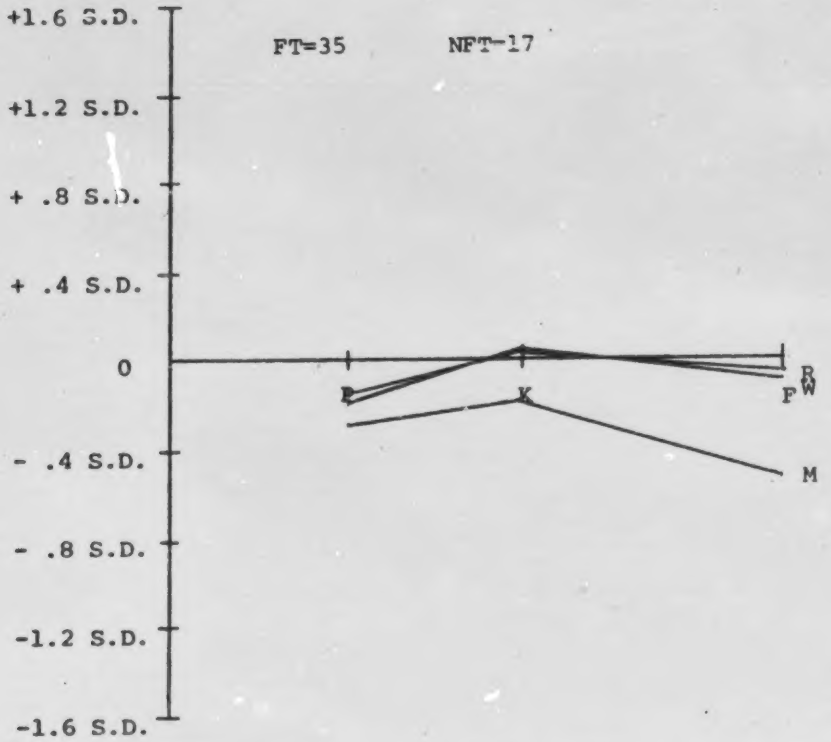
1.6.2 Effects by Site

The site level cumulative (Figure I-20 to I-25) show substantial variability. The profiles do, however, distinguish between the sites where the FT and NFT groups are closely matched and those where they are poorly matched. The profiles for the sites with closely matched FT and NFT groups, shown in Figures I-20 through I-23, display no significant differences between FT and NFT across the two-year period. The FT and NFT groups in these sites are progressing at the same rate.

The two sites where the FT and NFT groups are severely mismatched, shown in Figures I-24 and I-25, show large and significant losses in the FT group.

1.6.3 Summary

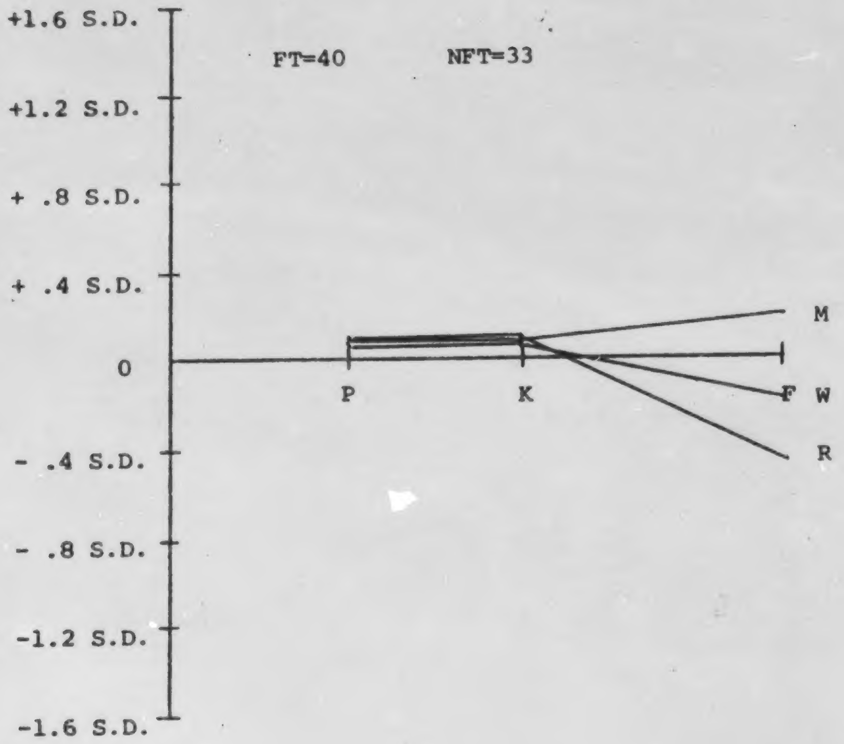
In summary, the High Scope Program has some positive influences. The effects of the program do, however, vary substantially between the more closely matched and the poorly matched sites. The fact that the FT children in the more closely matched sites appear to be keeping up with the NFT children suggests that this "cognitively oriented" program produces effects which are measurable on standard achievement tests. The severe losses during first grade in the poorly matched sights are difficult to explain with the available data. These losses may be due to specific factors such as implementation or social stratification factors which



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .0186	- 5.2639	- .7445
S.E.	.0606	2.7940	4.0690
Two Year Difference (B)	.0179	- 2.3536	1.4277
S.E.	.0542	2.6025	3.4808
First Grade Difference (C)	- .0294	- 3.5924	- 1.3326
S.E.	.0519	2.6748	3.3960
Outcome S.D.	.2033	9.6843	13.6330
P=(A-B)/S.D.	- .1795	- .3005	- .1590
K=(A-C)/S.D.	.0531	- .1726	.0431
F=A/S.D.	- .0915	- .5435	- .0546

Figure I-20 First Grade Outcome Profile for High Scope Site NE

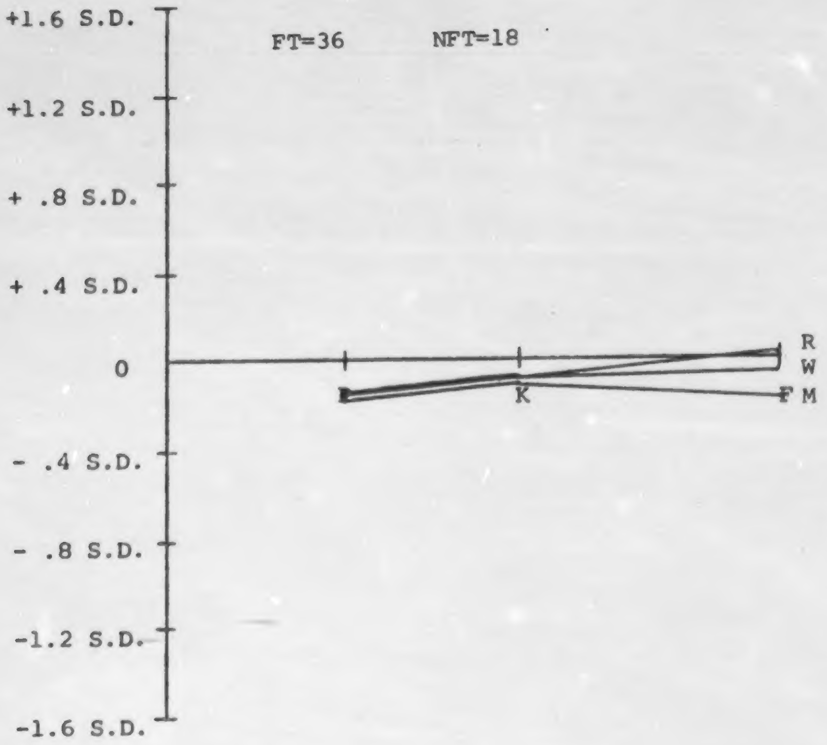
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	-.038	1.863	-6.354*
S.E.	.050	2.281	2.543
Two Year Difference (B)	-.048	1.291	-7.230*
S.E.	.046	1.836	2.078
First Grade Difference (C)	-.048	1.273	-7.246*
S.E.	.042	1.744	2.022
Outcome S.D.	.213	9.679	11.200
P=(A-B)/S.D.	.047	.059	.078
K=(A-C)/S.D.	.047	.061	.080
F=A/S.D.	-.178	.192	- .567

Figure I-21 First Grade Outcome Profile for High Scope Site NC

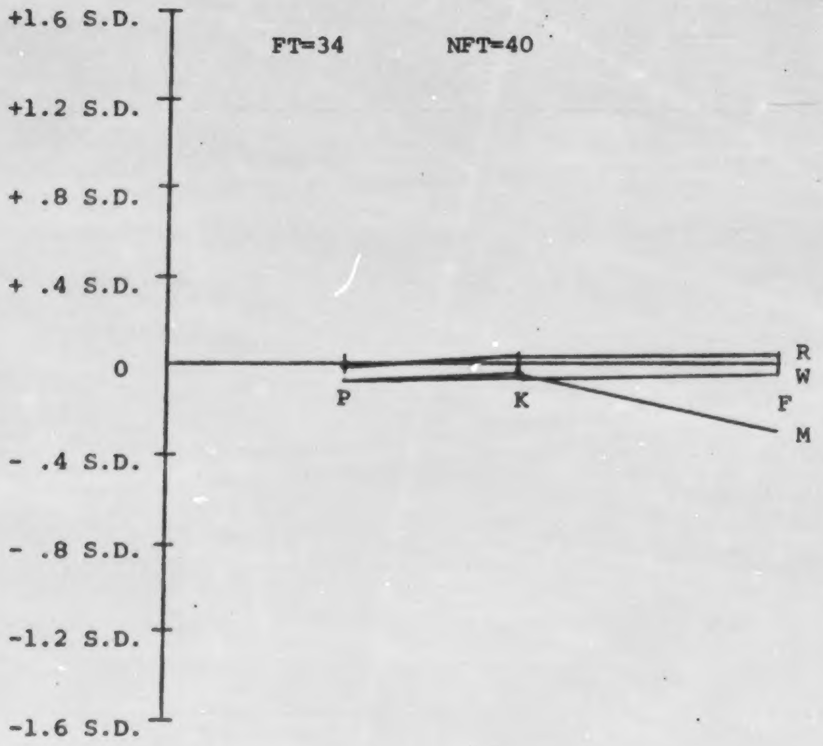
Note: All items marked by * are significant (p ≤ .05).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .0123	- 1.8889	.2778
S.E.	.0581	3.0477	4.0671
Two Year Difference (B)	.0194	- .1389	2.2164
S.E.	.0499	2.4821	3.8019
First Grade Difference (C)	.0021	- .8129	1.2969
S.E.	.0410	2.2370	3.5195
Outcome S.D.	.1996	10.4960	13.9559
P=(A-B)/S.D.	- .1588	- .1667	- .1389
K=(A-C)/S.D.	- .0721	- .1025	- .0730
F=A/S.D.	- .0616	- .1800	.0199

Figure I-22 First Grade Outcome Profile for High Scope Site W1

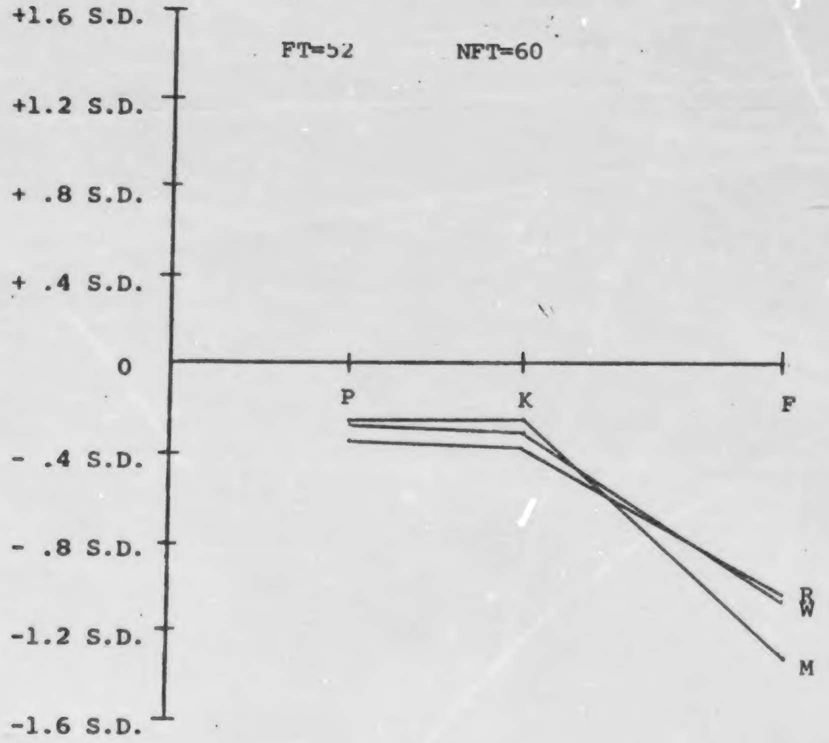
Note: All items marked by * are significant (p ≤ .05)



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0081	- 3.2279	1.1691
S.E.	.0500	2.4796	3.0183
Two Year Difference (B)	.0095	- 2.4728	2.0932
S.E.	.0467	2.3510	2.4683
First Grade Difference (C)	.0017	- 2.7596	1.8470
S.E.	.0406	2.2024	2.3721
Outcome S.D.	.2129	10.6807	12.8639
P=(A-B)/S.D.	- .0065	- .0706	- .0718
K=(A-C)/S.D.	.0300	- .0438	- .0526
F=A/S.D.	.0380	- .3022	.0908

Figure I-23 First Grade Outcome Profile for High Scope Site W2

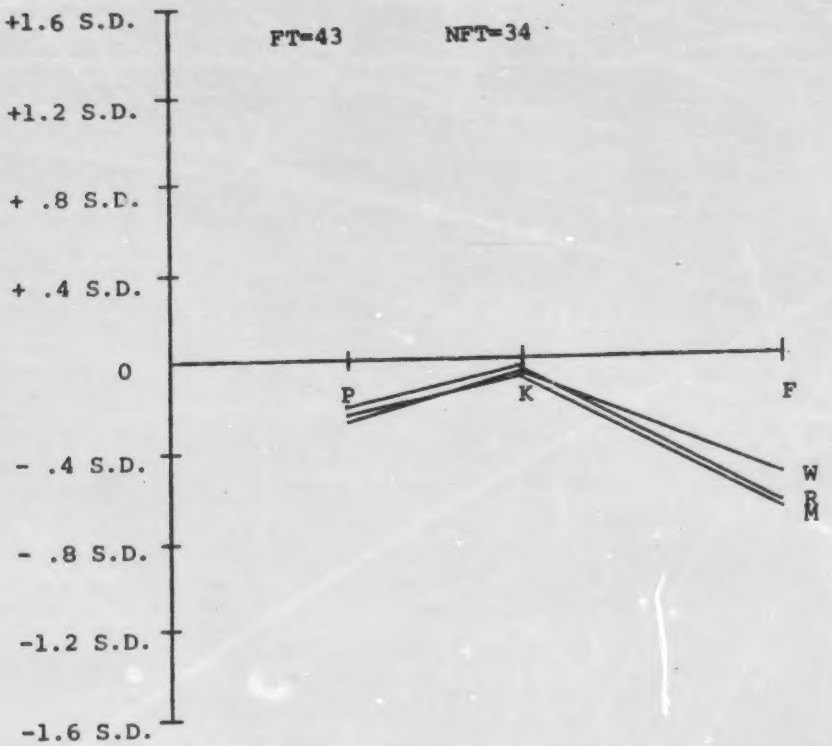
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	-.2689*	-16.3995	-15.9689*
S.E.	.0469	2.1272	3.0118
Two Year Difference (B)	-.2013*	-13.2037*	-10.7997
S.E.	.0393	1.6930	2.3676
First Grade Difference (C)	-.1928	-13.3775*	-10.3405*
S.E.	.0387	1.7061	2.3470
Outcome S.D.	.2378	12.1094	14.9086
P=(A-B)/S.D.	-.2843	-.2639	-.3467
K=(A-C)/S.D.	-.3200	-.2496	-.3775
F=A/S.D.	-1.1308	-1.3543	-1.0711

Figure I-24 First Grade Outcome Profile for High Scope Site S

Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .1352*	- 8.9737*	-11.3859
S.E.	.0578	2.8320	3.6857
Two Year Difference (B)	- .0675	- 5.7855*	- 7.4850*
S.E.	.0461	2.3370	3.0960
First Grade Difference (C)	- .1202*	- 8.0320*	-10.8264*
S.E.	.0426	2.2459	2.9011
Outcome S.D.	.2546	12.8326	16.6505
P=(A-B)/S.D.	- .2659	- .2484	- .2342
K=(A-C)/S.D.	- .0589	- .0733	- .0336
F=A/S.D.	- .5310	- .6992	- .6838

Figure I-25 First Grade Outcome Profile for High Scope Site W3

Note: All items marked by * are significant ($p \leq .05$).

act as impediments to a child's development. The losses may also be related to the inappropriateness of the program for this target group of children. The explanation of the differences between these sites must wait for additional data from later grades and for the exploration of variables that define the differences among sites and their children. Since High Scope is a child-centered program emphasizing individual differences, the variability across site may be related to the children's readiness. Thus, the program may be developing important skills, which are not reflected on achievement tests until the children have reached a threshold level of problem solving skill and confidence. Additional years of data will confirm or disconfirm this hypothesis.

1.7 University of Florida Parent Education Model

The University of Florida program is described as a Parent Education Model. Based on the premise that children's learning takes place as much at home as in school, the major objective is to improve children's school achievement through educating parents to participate directly in the education of their children. While the curriculum is not standardized, it does have Piagetian orientation.

This evaluation of the Florida program examines data from two types of sites. The first consists of three communities in the South and Far West. In these communities, though the FT and NFT groups differ demographically, they are roughly equal on kindergarten entry level tests. All of these sites display a high degree of social stratification by either ethnic membership or socioeconomic status and the FT group, to a lesser extent, come from the lower strata.

The second category consists of a small North Central city in which the FT/NFT children are almost exclusively White. Though both groups have a rather low SES, they are not as poor as the families in the first site category. The FT children in this site score lower on entry level tests than the NFT children. The lower entry test scores, in the absence of SES differences, suggest that these children are disadvantaged in a manner not directly reflected in SES--but perahps indicative of an educational disadvantage or learning disability.

1.7.1 Site Descriptions

The first site category contains two large Southern cities, one in the Southeast (S1), the other in the Southwest (S2), and a small town in the Far West (W).

In the large Southeastern city (S1), 86 percent of the FT children are Black, compared with 50 percent in the NFT group. The FT children come from families with significantly lower incomes than those in the NFT group. (See Table VI-26, Vol. II-A).

In the Southwestern city (S2), virtually all the FT/NFT children are from minority groups. Of the FT children, 63 percent are Black and 35 percent Spanish-surnamed; of the NFT children, 42 percent are

Black and 55 percent Spanish-surnamed. Both groups have very low SES as shown in Table VI-27, Vol. II-A.

In the small Far Western town (W), both FT and NFT children are predominantly White, but the families of FT children have substantially lower incomes than NFT families. (See Table VI-28, Vol. II-A.)

At all three sites there are not differences between the two groups on kindergarten entry level tests. The FT children score as high as or higher than the NFT children.

In NC1, the second type of site, both FT and NFT groups are White and essentially equivalent on measures of parental education, family income and family structure. In this small North Central community, both groups appear to come from intact homes with somewhat higher SES than the FT families in the first site category. However, the FT children's kindergarten entry level test scores are lower than those of the NFT children. (See Table VI-29, Vol. II-A.)

1.7.2 Effects by Site

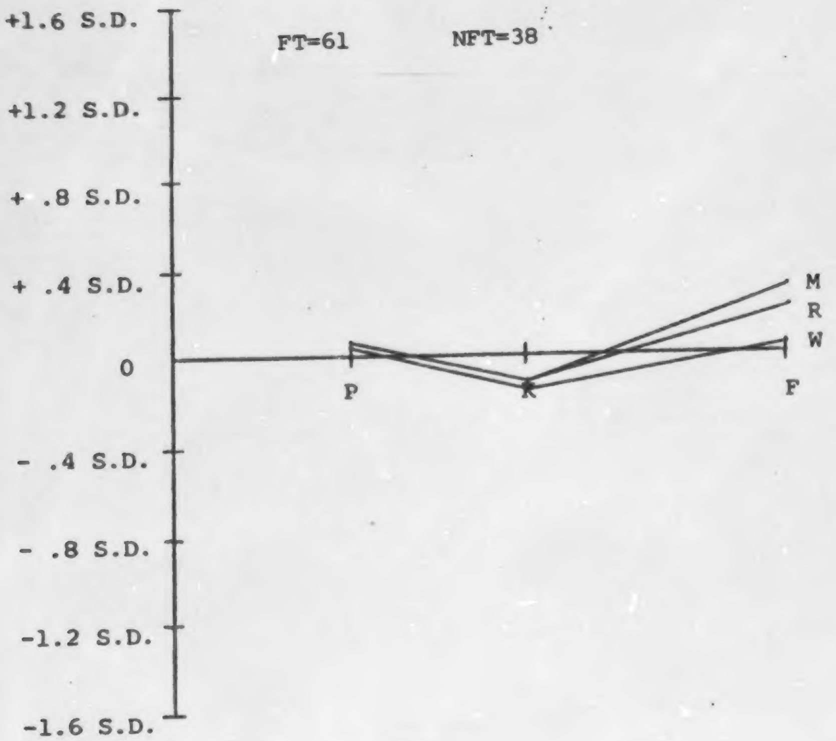
At the first three sites (S1, S2, W), the Ft children show gains in some skill areas during first grade (Figures I-26 to I-28). There are positive trends in Reading and Word Analysis at all three sites. Though these results are not statistically significant, the consistency across site suggests a stable, though small, pattern of growth over time.

The Mathematics findings are not consistent across these sites. At one site (S1) the program has a positive first grade effect, at another (W) the program has no significant effect, and in the third (S2) the FT group shows a loss in first grade.

In site NC, a small town with somewhat higher SES than the other sites but where FT have lower entry scores, the results are rather different. FT children at this site demonstrate a cumulative loss in all outcomes with some tendency to maintain their position on Word Analysis in first grade. (See Figure I-29.)

1.7.3 Summary

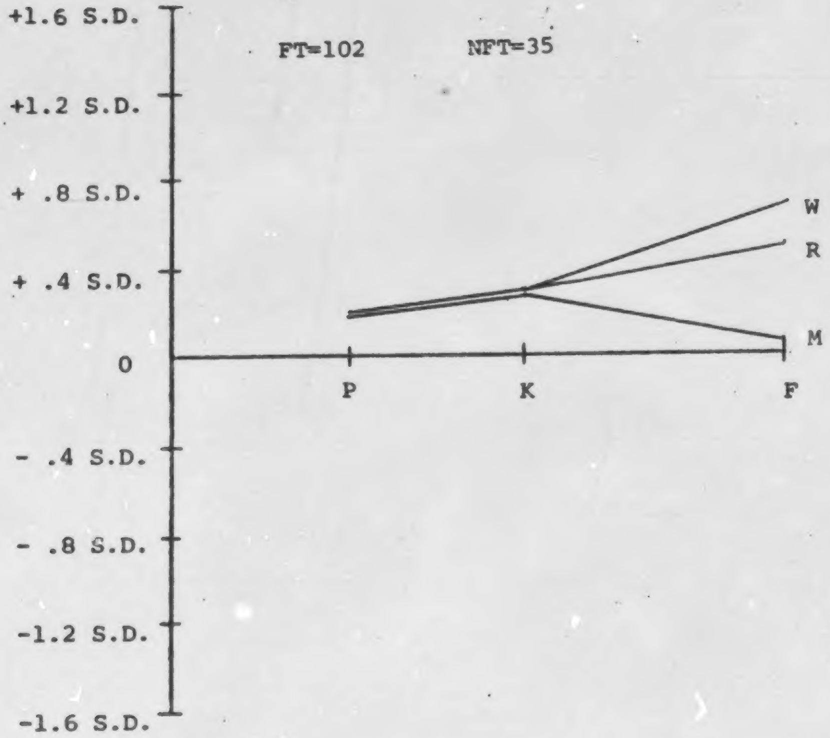
It is clear that the Florida program is, at the end of first grade, moderately successful in some skill areas at three geographically



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0114	3.4072	3.2140
S.E.	.0537	2.2584	3.0854
Two Year Difference (B)	.0024	2.7501	2.3620
S.E.	.0497	1.9903	2.8475
First Grade Difference (C)	.0474	4.7383*	5.0239
S.E.	.0448	1.7376	2.5436
Outcome S.D.	.2585	10.9990	14.9359
P=(A-B)/S.D.	.0348	.0597	.0570
K=(A-C)/S.D.	- .1392	- .1210	- .1211
F=A/S.D.	.0441	.3097	.2151

Figure I-26 First Grade Outcome Profile for Florida Site S1

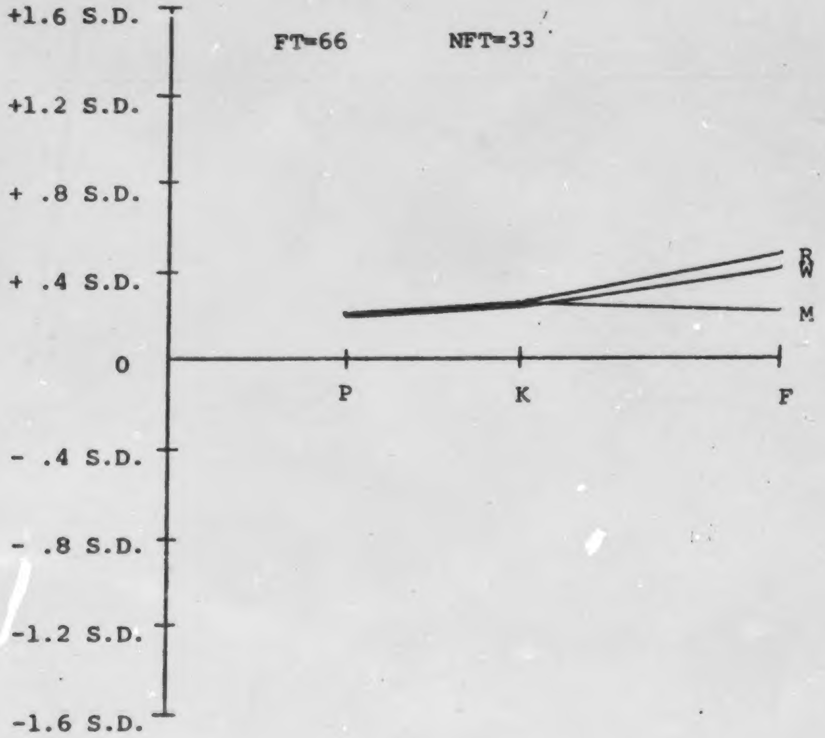
Note: All items marked by * are significant (p ≤ .05).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.1674*	.6860	6.7647*
S.E.	.0440	2.1634	2.5611
Two Year Difference (B)	.1200*	- 1.1616	4.0252
S.E.	.0378	2.0305	2.2479
First Grade Difference (C)	.0958*	- 2.1804	2.6983
S.E.	.0316	1.8425	1.9454
Outcome S.D.	.2354	11.0071	13.3579
P=(A-B)/S.D.	.2013	.1678	.2050
K=(A-C)/S.D.	.3041	.2604	.3044
F=A/S.D.	.7111	.0623	.5064

Figure I-27 First Grade Outcome Profile for Florida Site S2

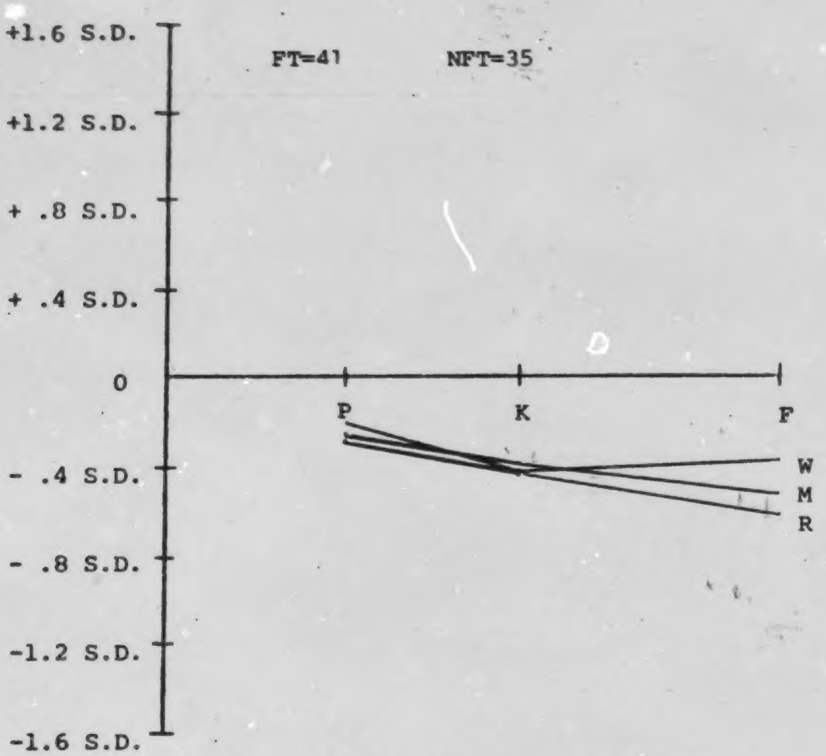
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.1073	2.4705	8.2973*
S.E.	.0549	2.3946	3.5836
Two Year Difference (B)	.0576	.1231	4.7627
S.E.	.0467	2.0003	2.9972
First Grade Difference (C)	.0483	- .1749	4.2183
S.E.	.0397	1.8432	2.6292
Outcome S.D.	.2552	10.9746	16.7906
P=(A-B)/S.D.	.1947	.2138	.2105
K=(A-C)/S.D.	.2311	.2410	.2429
F=A/S.D.	.4204	.2251	.4941

Figure I-28 First Grade Outcome Profile for Florida Site W

Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .0947	- 7.1958*	-10.5561*
S.E.	.0574	2.9610	3.7338
Two Year Difference (B)	- .0405	- 3.8039	- 5.7714*
S.E.	.0490	2.2225	2.4996
First Grade Difference (C)	.0133	- 1.8138	- 3.2356
S.E.	.0456	2.1584	2.3713
Outcome S.D.	.2522	13.2804	16.9639
P=(A-B)/S.D.	- .2149	- .2554	- .2820
K=(A-C)/S.D.	- .4282	- .4052	- .4315
F=A/S.D.	- .3754	- .5418	- .6222

Figure I-29 First Grade Outcome Profile for Florida Site NC

Note: All items marked by * are significant ($p \leq .05$).

scattered sites. The FT children at these sites come from Black and/or poor families, but there is no evidence that there is educational disadvantage when they enter kindergarten. Promoting parental involvement in children's home-learning activities thus seems to be an effective input in this kind of a situation. On the other hand, where the FT children come from families whose SES is similar to the NFT families, but where these children entered kindergarten at a much lower achievement level than the NFT children, the parent involvement model was unable to keep the FT children from falling further behind over the years.

In order to understand the Florida model, it is necessary that extended analysis be undertaken. The patterns need to be replicated in a broader range of sites and over a more extended grade span before conclusions about the impact of a parent training model of this kind can be drawn. This is particularly true in the light of the apparently greater effects in first grade than in kindergarten, a pattern which might be uniquely expected in a parent training model. By definition, the effects on children must appear sometime after the training takes place since the educational treatment involved is the relationship between the trained parent and the child. It can be expected that if parent training starts when the child enters kindergarten, that the "treatment" would start at a later date and that the effects will emerge still later. The more difficult sites (e.g., site NC) must be expected to be slower in showing effects in this kind of a model. The present findings are suggestive of this process, but further data are required before conclusions can be drawn.

1.8 The Educational Development Center (EDC)

The EDC program is based on the belief that learning is facilitated by a child's active participation in the learning process. In the flexible, open classroom environment, children are encouraged to initiate activities, pursue their interests, and generally assume responsibility for their own learning. The basic objective of the program is to provide the optimal environment for children's growth in academic and problem solving skills, self-expression and self-direction.

This evaluation of the EDC program examines data from three sites, all large cities in the Northeast and Middle Atlantic areas. The FT groups at all sites have lower kindergarten entry scores and lower SES than their NFT counterparts.

1.8.1 Site Descriptions

At site NE1, a small Northeastern city, both FT and NFT children are White, but the FT children have lower SES and kindergarten entry scores than the NFT comparison group. A considerably higher percentage of FT children attended preschool (FT = 60%; NFT - 10%). (See Table VI-30, Vol. II-A.) Nevertheless, substantially more Title I money is expended in the NFT schools.

At site NE2, a large Northeastern city, both FT and NFT groups are predominantly Black (FT = 98% Black; NFT = 82% Black, 15% Spanish-surnamed). FT and NFT are similar on measures of SES, except that FT family income is lower. More FT children than NFT children attended preschool (FT = 87%; NFT = 21%); however, the two groups show no differences on kindergarten entry level tests. (See Table VI-31, Vol. II-A.) One surprising difference between the groups is that the NFT first grade classes are considerably smaller, 16 children compared with 26 children in the FT classes. This contrast threatens the validity of any comparison of these groups since class size was not used as an adjusting factor in our analyses.

The remaining EDC site (S) is a large Middle Atlantic city. Both FT and NFT groups are predominantly Black and very similar on all SES indicators. However, the FT group is a full standard deviation behind the NFT group on the kindergarten entry WRAT. In addition, 100 percent of the FT children and none of the NFT children have had some form of preschool experience. (See Table VI-32, Vol. II-A.) In this site, as in the others, the FT group displays a disadvantage which is somewhat difficult to specify.

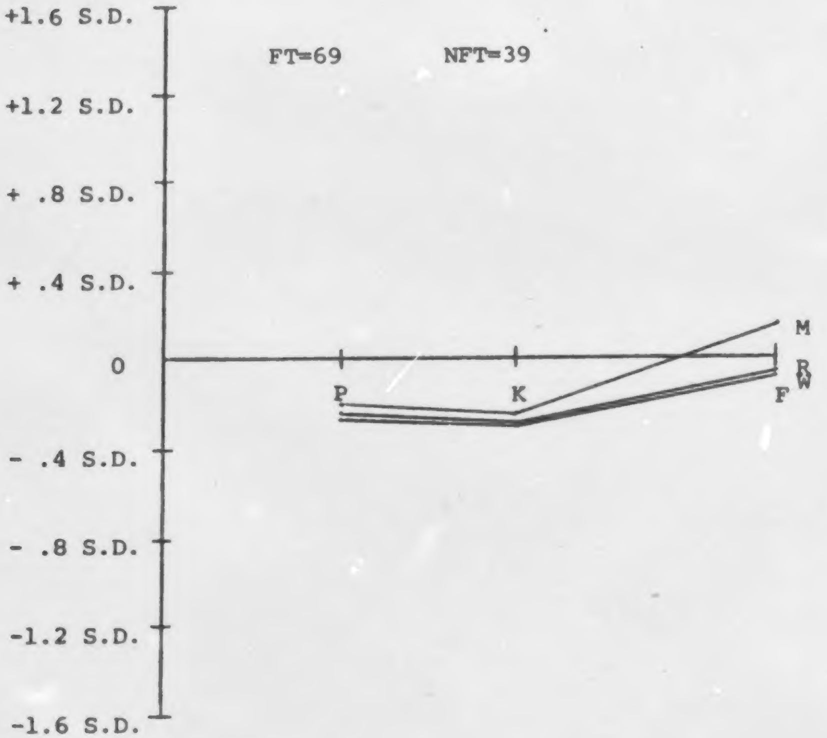
1.8.2 Effects by Site

The first grade cumulative profiles for the EDC sites are shown in Figures I-30 through I-32 The figures indicate some variability among sites in kindergarten, with more consistency in first grade. The FT children in all sites show substantial gains in Mathematics during first grade. Furthermore, the FT children in two sites, sites S and NE1, show first grade gains in Reading and Word Analysis as well. Only in site NE2, the site with exceptionally small NFT classes in first grade, do the NFT children gain more than the FT children in Word Analysis and Reading. In site S, the FT children show a cumulative gain over the kindergarten to first grade interval, while in the other sites, the gain occurs exclusively in first grade.

1.8.3 Summary

The site analysis suggests that the EDC program consistently produces gains on first grade outcome measures in Math, and in two out of three sites, in Reading and Word Analysis. The stability of the effects across such a variety of sites and children suggests that at least in the Eastern region of the country, the open-classroom program does have an impact on skills measured in the standard achievement mode.

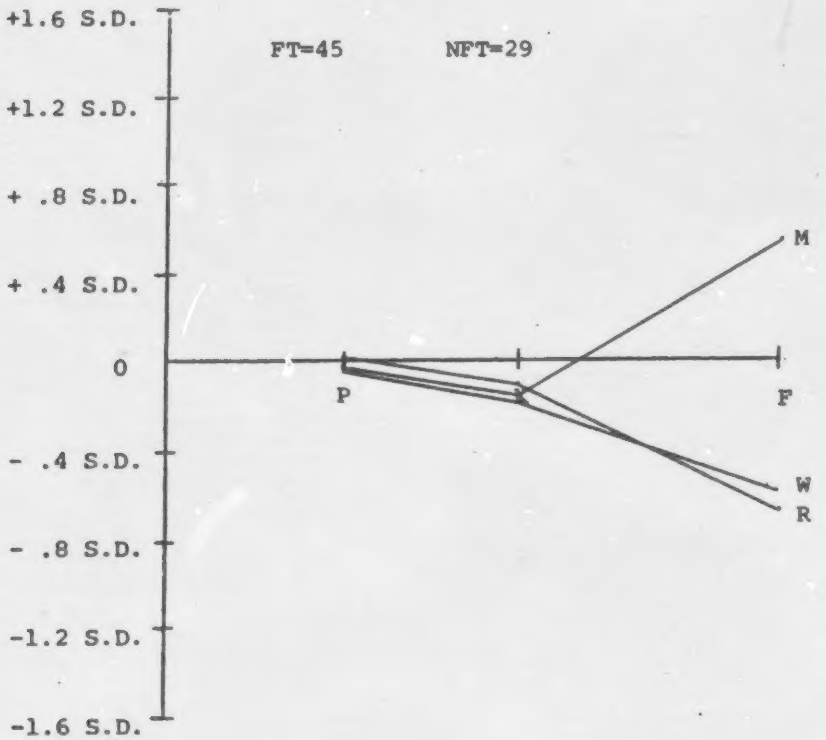
Although the generally positive first grade effects are found in a variety of sites, the one factor which differentiates FT from NFT is the much larger number of FT children who attended preschool. The consequences of this fact on kindergarten entry scores are negligible as they are for the end of kindergarten effects. By the end of first grade, however, the EDC children are showing particular strength in math problem solving. It is reasonable to conclude that the EDC program was able during the kindergarten year to socialize children and teachers into the open, discovery-oriented system in which methods of manipulating



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .0186	1.6555	- .9487
S.E.	.0497	2.1237	3.3526
Two Year Difference (B)	.0492	3.9252*	3.2484
S.E.	.0426	1.8975	2.8848
First Grade Difference (C)	.0587	4.2800*	3.8589
S.E.	.0361	1.6790	2.4460
Outcome S.D.	.2471	10.5813	16.6630
P=(A-B)/S.D.	- .2743	- .2145	- .2518
K=(A-C)/S.D.	- .3128	- .2480	- .2885
F=A/S.D.	- .0752	.1564	- .0567

Figure I-30 First Grade Outcome Profile for EDC Site NE1

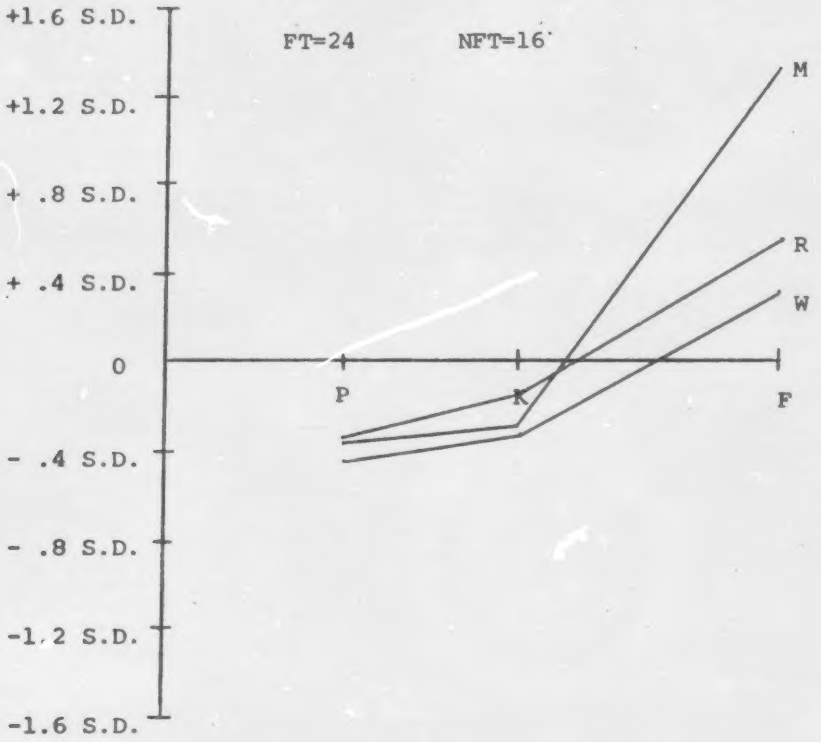
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .1405*	5.8521*	- 9.2628*
S.E.	.0543	2.4502	3.0791
Two Year Difference (B)	- .1284*	6.1595*	- 9.3789*
S.E.	.0515	2.2507	2.7183
First Grade Difference (C)	- .0966*	7.5038*	- 7.8331*
S.E.	.0476	2.0931	2.5526
Outcome S.D.	.2369	10.6157	13.6244
P=(A-B)/S.D.	- .0510	- .0289	.0085
K=(A-C)/S.D.	- .1853	- .1555	- .1049
F=A/S.D.	- .5930	.5512	- .6798

Figure I-31 7 First Grade Outcome Profile for
EDC Site NE2

Note: All items marked bv * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0691	18.9615*	8.1978
S.E.	.0774	3.7145	5.1075
Two Year Difference (B)	.1704*	24.1393*	13.7553*
S.E.	.0686	3.1372	4.5280
First Grade Difference (C)	.1430	23.1618*	10.5245
S.E.	.0709	3.2765	4.3873
Outcome S.D.	.2267	14.1347	15.3235
P=(A-B)/S.D.	- .4468	- .3663	- .3626
K=(A-C)/S.D.	- .3259	- .2971	- .1518
F=A/S.D.	.3048	1.3414	.5349

Figure I-32 First Grade Outcome Profile for EDC Site S

Note: All items marked by * are significant ($p \leq .05$).

the materials of growth is emphasized rather than the specific skills of word and number analysis. Little of this generalized to performance on standardized kindergarten tests. The more mature first graders, after a year of preschool and a year in this kind of kindergarten are now able to translate the number games and numerical-concept experiences into successful solutions on the math concept problems which are found in the first grade test. An examination of the specific strenghts of these children and a test of the dependence of these strengths on either one or two years of socialization into this educational mode is required in order to advance our understanding of these EDC effects. Specifically, we need to know if these, and future effects of the EDC program are dependent upon preschool experiences of particular kinds, or whether these effects can be generated within the traditional public school mode.

1.9 University of Pittsburgh, Primary Education Project

The Primary Education Project utilizes a number of interrelated curriculum components which are carefully structured and sequenced to provide for optimally efficient learning. Three general classes of skills are included in these curriculum components which are designed to form the foundation of all higher level functioning: (1) orienting and attending skills, (2) perceptual motor skills, and (3) conceptual-linguistic skills. These latter include classification, reasoning, memory, language, and mathematics concepts. The curriculum is highly individualized in order to allow the child to progress at his own pace. The teacher serves as a facilitator and resource person as the child moves through each component.

Only two out of eight Pittsburgh sites met the criteria for inclusion in these analyses: a small Northeastern city and a rural community in the North Central region of the country. The assessment of the effectiveness of a program based on only two sites has inherent weaknesses. Furthermore, the two sites in which there are sufficient numbers of tested children differ substantially from one another on a variety of pupil characteristics as well as on the nature of the FT/NFT match within site. These differences prevent examination of Sponsor effects independent of specific site or child variables. Thus the findings for this Sponsor must be interpreted with caution.

1.9.1 Site Descriptions

The FT group in the North Central community (NC) is one of the highest in the analytic sample with respect to family income, level of parental education, and occupational status. Most of the children are White, come from intact families, and have attended preschool programs. The NFT group in this site is similar in both SES and other characteristics; moreover, the two groups are well matched on entry level tests.

The FT group in the Northeastern site (NE) is substantially lower than the NFT group on measures of SES, has a somewhat lower percentage of intact families, and has substantially lower pretest scores. (See Table VI-34, Vol. II-A.)

1.9.2 Effects by Site

The site level cumulative profiles are shown in Figures I-23 and I-34. The FT children in the North Central site are holding their own in Mathematics and Reading at the end of the two-year period. In Word Analysis, the FT children appear to be making slightly greater progress than the NFT children during first grade.

On the other hand, FT children in the Northeastern site appear to be losing ground between kindergarten entry and the end of first grade. In Reading and Word Analysis, the cumulative deficit is substantial; in Mathematics, it is minimal.

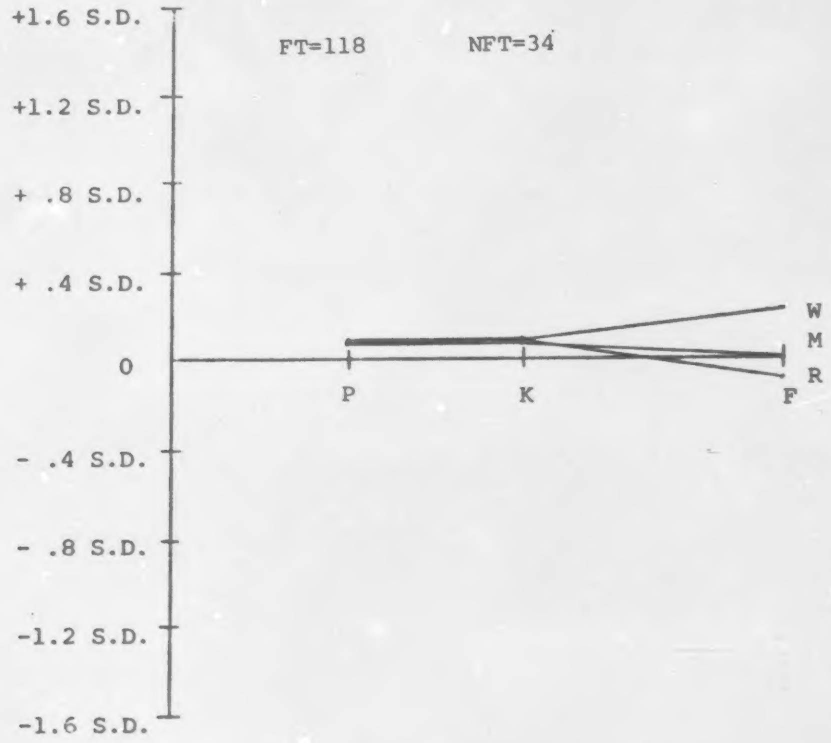
1.9.3 Summary

The use of caution in the interpretation of the Pittsburgh model results cannot be overemphasized. Only two sites are represented in these findings, one characterized by an unusually high FT/NFT group in terms of both SES and entering achievement, and the other by a severe FT/NFT mismatch.

The program does not, however, appear to be significantly affecting children's performance in either mathematics or the other achievement domains at the end of first grade. In one site, where both FT and NFT children are relatively high on entering achievement, they remain roughly equal at the end of first grade. In the other site, where the FT children enter school behind their NFT counterparts, the gap between them continues to widen.

Clearly the first grade tests are not sensitive to the specific objectives of the Pittsburgh curriculum during the second year. This model, which appears to be appropriate for kindergarten, at least in the math area, does need reshaping for first grade. Without additional data from both other sites and later grades, however, little can be said of this Sponsor's effectiveness with certainty.

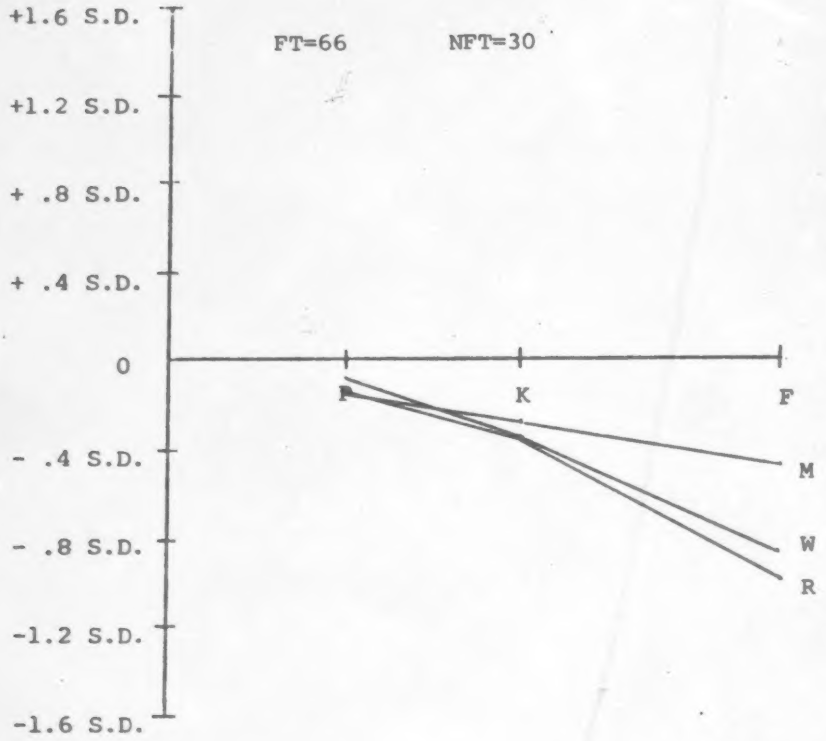
If we assume that in the absence of an experimental treatment, the high entry level children should remain at the same level as their well matched comparisons, and the low entry level children should fall further and further behind their "average" level comparisons, the Pittsburgh model appears not to meaningfully influence the expected progress in these two sites. The FT children perform as if there was no experimental treatment, except in kindergarten math.



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0317	.1142	- 1.0075
S.E.	.0271	1.7924	2.4950
Two Year Difference (B)	.0189	- .4945	- 2.0731
S.E.	.0235	1.6181	2.0270
First Grade Difference (C)	.0175	- .5629	- 2.1887
S.E.	.0224	1.5808	1.9349
Outcome S.D.	.1392	9.1780	12.7827
P=(A-B)/S.D.	.0919	.0663	.0833
K=(A-C)/S.D.	.1020	.0737	.0924
F=A/S.D.	.2277	.0124	- .0788

Figure I-33 First Grade Outcome Profile for Pittsburgh Site NC

Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .2131*	- 5.2576*	-16.2000*
S.E.	.0484	2.4061	3.1591
Two Year Difference (B)	- .1953*	- 3.5347	-14.0044*
S.E.	.0416	1.8441	2.6453
First Grade Difference (C)	- .1294*	- 2.1540	-10.3807*
S.E.	.0378	1.8929	2.5115
Outcome S.D.	.2399	11.1423	16.1444
P=(A-B)/S.D.	- .0741	- .1546	- .1359
K=(A-C)/S.D.	- .3488	- .2785	- .3604
F=A/S.D.	- .8882	- .4718	- 1.0034

Figure I-34 First Grade Outcome Profile for Pittsburgh Site NE

Note: All items marked by * are significant ($p \leq .05$).

1.10 Southwest Education Development Laboratory (SEDL)

The Language Development (Bilingual and/or English-as-a-second language) Approach was originally designed as an instructional program for predominantly Spanish-speaking classrooms. The primary emphasis of the approach is on language development; language is seen as the tool for acquiring a variety of skills including non-linguistic skills. Building upon the child's native language and culture, the kindergarten program stresses the development of visual, auditory, and motor skills, as well as thinking, discovery, and English language structures.

This evaluation of the SEDL program analyzes data from three sites -- one in a large Northeastern city (NE), the other two in small towns in the South (S) and West (W). Each of these communities has a substantial non-English-speaking population. Two of them (NE and W) contain a large number of children whose primary language is Spanish; at the remaining site (s), for a large number of children, Cajun French is the primary language. The FT and NFT children within these sites differ somewhat in demographic characteristics.

1.10.1 Site Descriptions

The characteristic shared by all of the SEDL sites is their substantial non-English-speaking population. Two of them (NE and W) have a large Spanish-speaking population, while the remaining site (S) has a large number of people whose primary language is Cajun French. Aside from this basic similarity, the sites differ in several important respects.

At site NE, a large city in the Northeast, the FT group consists of inner city Black and Puerto Rican children. At this site, much of the parent interview data are missing, making it impossible to compare the SES and family structure characteristics of the FT and NFT groups. The FT group has a higher percentage of children with preschool experience (FT = 86%; NFT = 30%) and shows no differences on the kindergarten entry level test. (See Table VI-35,Vol. II-A.

At site W, 82 percent of the FT children have Spanish surnames compared with 33 percent of the NFT group. The FT group also has a lower level of family income than the NFT group, but a somewhat higher percentage of families with both a mother and a father present in the home (FT = 91%; NFT = 78%). These differences suggest that the FT group consists predominantely of poor Chicano Families which tend to be highly stable, and that the NFT children come from richer white families. (See Table VI-36, Vol. II-A.)

Site S differs from both other sites in that the FT and NFT groups are closely matched on SES and family structure characteristics. The FT group has a higher percentage of White pupils than NFT (48% versus 26%). The community from which both groups of children come has a high percentage of Cajun French speakers (82%) and is relatively low on SES indicators. (See Table VI-37, Vol. II-A.)

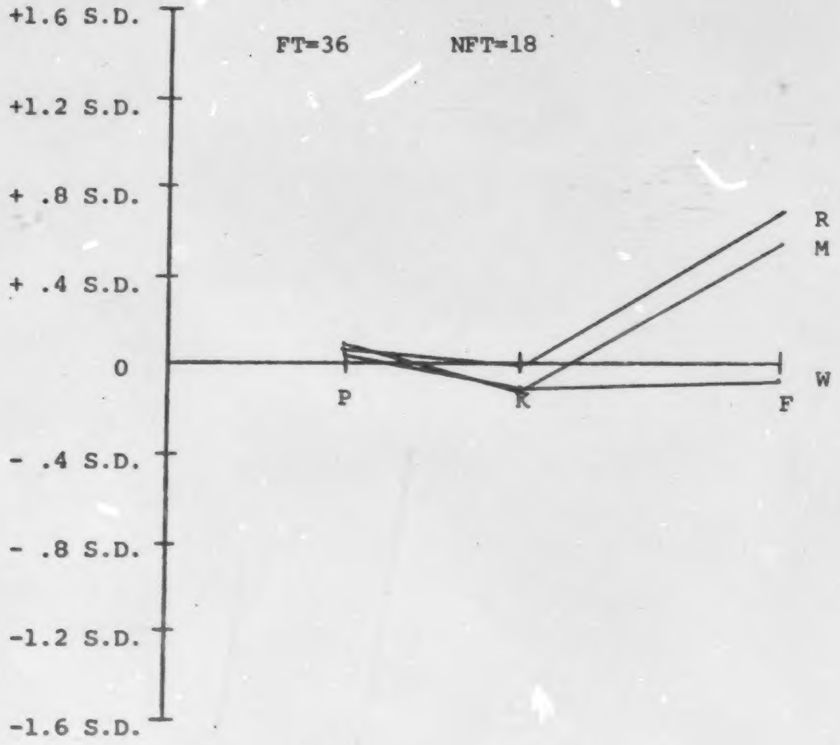
1.10.2 Effects by Site

The results for the SEDL sites shown in Figures I-33 through I-35, indicate substantial gains in Math at all sites and substantial gains in Reading at sites W and NE, both Spanish-language sites. Although the Work Analysis results are more variable, this finding is not incompatible with the phonological manipulation or word analysis.

1.10.3 Summary

The diversity of kindergarten effects for the SEDL program may be in part a function of the variable types of children with which it works, and in part a product differential implementation. Differences in both children and curriculum may be expected to produce maximum variability in the kindergarten year. The SEDL emphasis shifts in first grade to literacy skills and second language learning. The substantial first grade effects in Reading suggest that the program has succeeded in both of these phases.

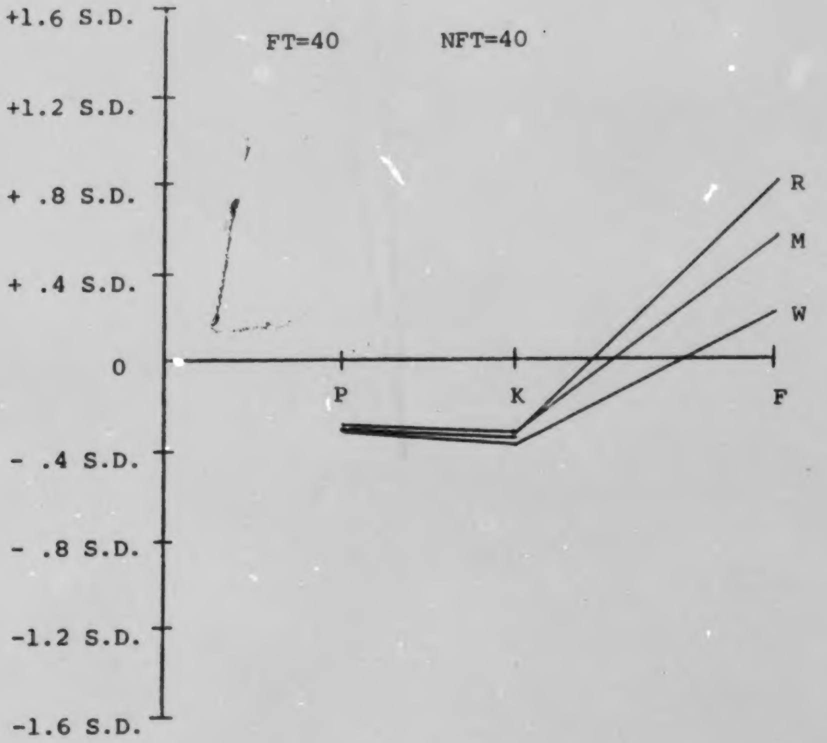
The less dramatic effects in the Southern site suggest that the SEDL program may be more appropriate for Spanish-speaking children than for the Cajun French. Given the confounding of children and site, however, it is impossible to say for sure whether the approach is simply inappropriate for this group or that it has not been as well applied.



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	- .0178	8.3849	9.8125*
S.E.	.0742	4.4550	3.9698
Two Year Difference (B)	- .0271	6.9708	8.9476*
S.E.	.0776	4.4886	4.1170
First Grade Difference (C)	.0081	10.0147*	10.0240*
S.E.	.0858	4.9096	4.5865
Outcome S.D.	.2466	15.3029	13.9480
P=(A-B)/S.D.	.0377	.0924	.0620
K=(A-C)/S.D.	- .1050	- .1065	- .0151
F=A/S.D.	- .0721	.5479	.7035

Figure I-35 First Grade Outcome Profile for SEDL Site NE

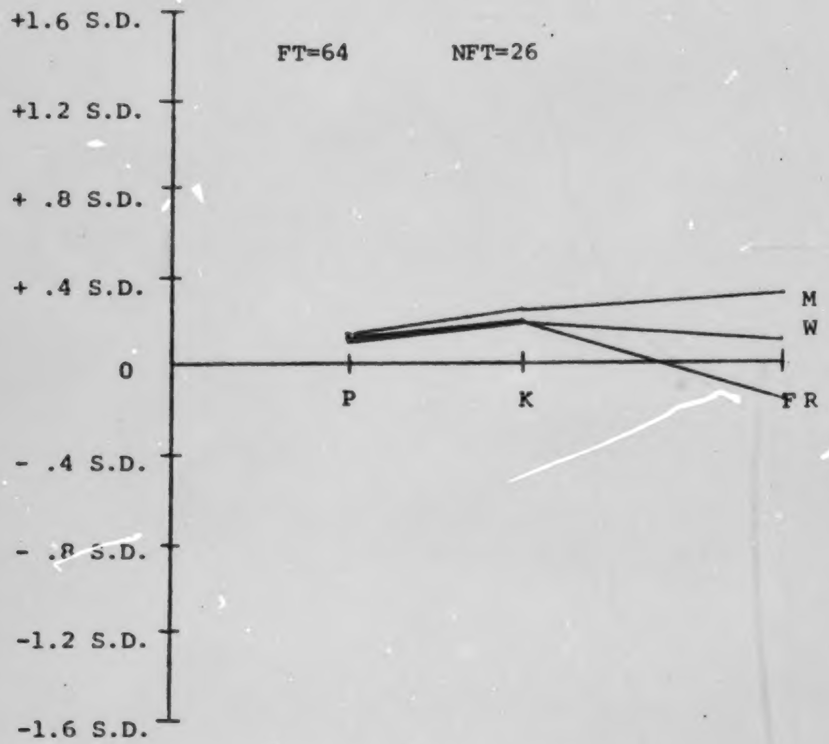
Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0389	5.9000*	9.5750*
S.E.	.0410	2.2676	2.4017
Two Year Difference (B)	.0998*	9.0961*	13.3258*
S.E.	.0355	1.8202	2.0624
First Grade Difference (C)	.1100*	9.4131*	13.6205*
S.E.	.0315	1.7589	2.0222
Outcome S.D.	.1832	10.5050	11.7095
P=(A-B)/S.D.	- .3324	- .3042	- .3203
K=(A-C)/S.D.	- .3881	- .3344	- .3454
F=A/S.D.	.2123	.5616	.8177

Figure I-36 First Grade Outcome Profile for SEDL Site W

Note: All items marked by * are significant ($p \leq .05$).



	MAT WORD ANALYSIS	MAT MATH	MAT READING
Unadjusted Difference (A)	.0217	4.0304	- 1.5353
S.E.	.0449	2.8036	2.2569
Two Year Difference (B)	.0027	2.4155	- 2.5646
S.E.	.0387	2.2678	1.9901
First Grade Difference (C)	- .0176	1.0391	- 3.4420
S.E.	.0369	2.1095	1.9357
Outcome S.D.	.1935	12.1971	9.7357
P=(A-B)/S.D.	.0981	.1324	.1057
K=(A-C)/S.D.	.2031	.2452	.1958
F=A/S.D.	.1121	.3304	- .1576

Figure I-37 First Grade Outcome Profile for SEDL Site S

Note: All items marked by * are significant (p ≤ .05).

1.11 Conclusions and Summary

There are three major points which need to be summarized:

- Generally, first grade effects are less impressive than kindergarten effects
- The effects of a specific program vary greatly from site to site
- It is still very difficult to relate findings to the educational model offered by Sponsors. The test of educational theories has not yet been accomplished in the Follow Through experiment.

Each of these points is discussed in turn below.

First, it is clear that the effects found at first grade are considerably less impressive than those found in kindergarten.* This is most apparent when the overall Sponsor effects are examined. If we review data from all available sites for each Sponsor to produce an overall Sponsor analysis, it is clear that only Kansas, EDC, and SEDL show gains of any consequence in first grade. Both EDC and Kansas show strong gains in mathematics, and SEDL shows such gains in both reading and math. The Kansas math gains are particularly interesting because they represent a continuation of the substantial gains produced by this Sponsor in these same children when they were in kindergarten. The effect is not completely uniform in all Kansas sites, as it is for the EDC and SEDL sites. In the one site where Kansas did not show gains in math, this Sponsor was dealing with a group of FT children who entered kindergarten with math scores that were very much higher than those of the NFT comparisons in that site. Despite this great advantage, the children in this site finished the kindergarten year no further ahead of the NFTs than they were at the beginning. This is quite a different picture from that found in the other Kansas sites, and suggests that the NFT group in this site received some special math treatment during the kindergarten year. This treatment may have continued during the first grade since the gap between the FT and the NFT narrowed by the end of the first grade. The point is, of course, that there are at least two ways of producing gains in this site: that which was supplied to the FT group by Kansas, and that which was supplied to the NFT group. Alternate routes to math success is obviously present, and the findings with respect to EDC and SEDL bear this out.

---

*Note should be taken that the estimates of first grade effects are based on a slightly longer time interval (Spring to Spring) than the time interval on which the kindergarten effects are based (Fall to Spring). The less impressive first grade effects might, therefore, have been influenced by the summer experiences between kindergarten and first grade, or by a negative effect of a lack of Fall testing in first grade. Data are not available to assess these factors.

These two Sponsors, both of whom failed to produce significant kindergarten effects in math, show very strong first grade effects. There is no doubt that Kansas, SEDL, and EDC have all adopted divergent techniques for math instruction, although we do not have measures of these variables. Nevertheless it is reasonable to assume that they are different from each other, although both EDC and Kansas are likely to have roughly similar counterparts among the remaining Sponsors. But it is only this small minority who show any effects at all, and the majority who show a reduced set of effects compared to the kindergarten year.

Such a reduction in effects from the kindergarten year could indicate a number of factors at work. Early effects in an innovative program often suggests that it is the notion of innovation rather than the content of the innovation which is producing the effects. As the excitement of the innovation dies down, so too would the effects die down. This is generally what has been found here. Except for the fact that these analyses were performed on data collected from Cohort III, the third group of children with which the Sponsors were involved, it would be reasonable to suggest the placebo hypothesis as one of the alternatives to account for the reduced first grade effects. It does not seem likely, however, that this hypothesis could account for the present findings, although a comparison of the kindergarten-first grade pattern for earlier Cohorts might throw some light on the issue. Such a comparison was excluded from the present report.

Another factor which could contribute to these findings is any difficulty Sponsors may have had in translating their theoretical models into practical first grade curricula. Another aspect of this factor might be that for some reason, it may be more difficult to administer innovative models in first grade than in kindergarten. Of course, if this were the case, then the Follow Through experiment would be about the implementation of models and not their educational efficacy. Each model could be more or less successful with children depending upon who attempted to translate the model into operational terms. It is likely that considerations of this kind are quite important and that the Follow Through experiment is not a test of a reinforcement model of education, or a cognitive stimulation model, or a parent education model. Rather, it is a test of the Kansas version of a token economy, or a High Scope version of cognitive stimulation, or a Florida version of a parent education model. Unfortunately, data to examine these theoretical models or even variations in any one version are not available because the Follow Through experiment was not designed to consider these questions. They remain critical but unmeasured and uncontrolled factors in the experiment.

A third factor has to do with the differences between the developmental needs of children in kindergarten and first grade. Very important shifts are taking place in the developmental process at this time. In the earlier stage socialization, rule learning, social controls, and group-based norms for behavior are the dominant non-cognitive concerns. Cognitively, the tasks are predominantly code cracking, i.e., the discovery that shapes have names and that there are rules that allow shapes to carry meanings. In first grade the tasks emphasize concept manipulation, problem solving, and the acquisition of a sense of personal competence. The skills underlying these activities develop at very different rates and are likely to be dependent upon the social context in which they develop. Thus, the family, the community, the relationship between the immediate community and the broader social structure, and the school as a social system, are all likely to have profound effects on the rate at which the cognitive and affective development of children takes place. It follows that very different rates of development of children are occurring in sites in which social structure differs and that older children show greater site specificity than younger children. This would suggest that, along with the implementation factors which are likely to vary a good deal from site to site, Sponsor effects ought to vary across sites. Site level analyses ought to reveal a number of important effects.

The second point which needs to be made here is that comparison among educational models is still very equivocal using these data. If we cluster Sponsors into any kind of theoretically coherent system, contradictions among the clusters begin to emerge. For example, the Oregon program shows kindergarten gains in reading and tends to maintain but not surpass this advantage in first grade. Their children's math scores however show sharp gains in kindergarten and precipitous losses relative to their NFT comparisons in first grade. The Kansas model, on the other hand, shows sharp gains in reading in kindergarten, which are generally not maintained in first grade, although the FT children are in some sites still ahead of the NFT children. However, the math effects are strong and consistent in both grades. The Oregon and Kansas effects are in almost the exact opposite direction. Bank Street and EDC, two Sponsors often clustered together on the basis of their common orientation toward a less structured, open classroom, cognitive model also show almost the opposite pattern of effects. Bank Street shows losses in reading generally, and in math in most sites. EDC, on the other hand, shows gains in both reading and math during the first grade. High Scope, Far West, and Arizona, all cognitive stimulation models, all show generally negative effects but so does Pittsburgh which structures the learning materials of first graders according to a very specific set of precisely reinforceable steps. The two programs whose models are not clearly set according to a particular pedagogical orientation (Florida, the parent education model, and SEDL, the bilingual model) both show generally positive effects, with SEDL showing very marked gains in math.

Despite the fact that clustering Sponsors according to some theoretical system produces inconsistent results, there still remains the fact that three Sponsors--Kansas, EDC, and SEDL--are producing significant effects at the end of first grade. After two years of working with Cohort III, those Sponsors which are going to be effective might begin to emerge. On the other hand, programs such as those developed by Oregon, Florida, High Scope, and Far West have shown a variety of significant effects in the past, and should be watched carefully in the future. There is still not enough known about the reasons for the successes and the failures of any of the Sponsors to be able to predict even with minimal certainty that the future will replicate the past. Nevertheless, the Sponsors about whom some discussion ought to be developed are Kansas, EDC, and SEDL.

Of these three, Kansas poses the greatest problems of interpretation. Clearly the token economy program, when implemented in large cities with Black children from low income families is a very powerful means of improving math achievement. This is true in both kindergarten and first grade which suggests that the effectiveness of this technique is not a function of the cognitive differences which differentiate these two levels of children. The framers of the Kansas program have been able to adapt the mechanics of the token reinforcement system to the kinds of materials which are central to first grade as well as kindergarten curricula. The nature of the token economy system is such that teachers are allowed some flexibility as to what is reinforced. Increasing the teachers' sensitivity to the moments when reinforcement is appropriate and to the techniques by which the reinforcement is maximally effective depends only in part upon a completely prescribed set of curriculum materials. The teachers, under this program, have several choices as to materials and certainly have a range of techniques in which they are trained, so that the transition from kindergarten to first grade appears to be smooth and effective.

The situation is quite different for those rural white children who are not particularly deficient in achievement upon entry into kindergarten. Here, the FT children show losses in first grade after a successful kindergarten year. One possible explanation for this loss is that the materials used in the first grade classrooms at this site may no longer be appropriate because these first graders are too advanced for them. Along with this, it is necessary to say that the materials still appear to be appropriate to the needs of the urban first graders for whom significant first grade effects emerge. If this hypothesis is true, the materials may cease working for the lower achieving children in the urban sites when these children reach the level already achieved by the rural children.

Another hypothesis is that the token economy program is not appropriate for the more advanced children of the rural site. The technique may not maintain the interest or commitment of the higher achieving rural children as it seems to for the lower achieving urban children.

There is no definitive way of examining either of these hypotheses but there are some data which enable us to explore the problem. First it is clear that the rural children are significantly ahead of the urban children at the end of the first grade on math scores (rural FT mean = 40.5; urban FT mean = 31.2). Thus, either the materials or the token economy technique may be inappropriate to the needs of these more advanced children. Second, the positive math effects in the urban sites are not replicated in reading. In other words, reading materials used in the urban sites are not very effective even when used by the same teachers who, using token economy techniques, are stimulating gains in math. These same materials are not at all successful with the rural children who are significantly ahead of the urban children. Very little of this advanced status in first grade reading or math can be attributed to the unique contribution of the Kansas program so that the most defensible hypothesis at the moment is that the curriculum materials rather than the reinforcement techniques appear to be limited in effectiveness for kindergarten and low achieving first graders. Of course, a strong interaction between reinforcement techniques and curriculum materials may exist, but a great deal of careful examination of on-site data is required to settle that issue. For the present it is possible to say that this Sponsor is very effective with Black urban children and that if these children follow the same course as some of the rural children with whom the Sponsor also works, these effects may diminish in the future. Since the effectiveness may have more to do with the materials than the method, it is entirely possible that with a careful reexamination of materials, any downward course in the urban sites may be averted.

We turn now to a consideration of the EDC effects. The findings for this version of the British Infant School model are a little less difficult to interpret than those for the previous program. The most impressive aspect of these findings is that very little gain was to be seen in the kindergarten results, and a great deal of relative gain is found at the end of first grade. The sites with which EDC is working are restricted in geographical locale to be sure, but within the Northeast to mid-Atlantic region, there is a good deal of variability. They include two middle-sized urban centers and one large city. Two of the three serve predominantly Black groups and one is a predominantly White group. In one site, the FT group is significantly below the comparison NFT in SES indicators, while the other two FT/NFT groups show no differences on these measures. However, in all sites, the vast majority of the FT children have had a full year of preschool experience before entering kindergarten whereas few of the NFT children have had such experiences. In all of these sites the FT children showed dramatic improvements relative to the NFT in math, and in two of the three sites they showed strong gains in reading as well. The reading findings were also true for word analysis, a test for which the EDC children were not directly prepared by the program.

EDC advisors (the name used by EDC for staff members who visit sites in order to facilitate the growth of the model on site) claim to have assumed the role of change agents interested in releasing the entire local school system from the confining structures of traditional organizational as well as pedagogical systems. Teachers' commitment to open school structures as well as open classrooms is perceived as the first step in providing an environment for children which would allow them to realize their potentials for growth. Clearly this kind of an approach requires a good deal of time and a great deal of patience in the face of major resistances in order to create the desired learning environment. The advisors have often noted that the British Infant model required a decade of development before it was so firmly accepted by teachers and community that it could operate without the handicap of being perceived as revolutionary. Assuming this to be true, the program would certainly be expected to show slow and uneven effects in the early years until the appropriate attitudes and values were established on the part of teachers, community, and children. At least with respect to the three sites included in these analyses, this pattern seems to be emerging. It is, of course, much too early to be definitive about this emergence since the pattern requires several more years to be verified. Further, the test of the hypothesis that appropriate acceptance of the revolutionary principles must precede real program effects cannot be made with the current data base. However, one important aspect of the program's expectations might be tested in the future.

It is clear that all of the positive effects are associated with children who have had preschool experiences before entering kindergarten. It may be that one year of a developmentally oriented preschool program, together with one year of an open classroom oriented kindergarten is the requisite condition for acquiring the open, discovery oriented fun-in-inquiry, approach to the learning environment that the EDCers wish to establish for children. One year in kindergarten may not be enough, and a year in a strongly developmentally oriented preschool program may be very helpful. Thus, the extent and nature of the preschool experiences of these children may represent an interesting opportunity to approximate a test of the "adaptation" hypothesis favored by this Sponsor.

Although this test has not yet been performed, there are some data to suggest support for the hypothesis. The site in which the most dramatic effects are found in all three domains is also the one in which most of the children attended a Head Start program which was run by EDC. There are some positive effects of this experience during the kindergarten year, but the first grade gains are dramatic indeed. It is also true that this site is a community controlled public school with a long history of innovation. These are confounding factors in one sense, and further evidence in support of the hypothesis in another sense. The excitement of innovation may be the source of the effects rather than the EDC model itself. On the other hand, the long tradition of innovation may be exactly the environment in which the EDC model flowers. An extensive examination of the EDC program in sites with other kinds of preschool programs, and with FT children who did not attend preschool is required to fully test these notions. For the present, it is sufficient to conclude that some of the potentials of this program may now be emerging and should be watched carefully in a variety of sites in the future.

We turn now to the SEDL effects. Here too a pattern has emerged of essentially no impact in kindergarten followed by reasonably sharp spurts in first grade. This is not easy to interpret because the pedagogy is not clear. Although originally intended to be a bilingual program, it is more likely an English as a Second Language program in some sites. The materials are still in process of development, but they are all geared to a population whose native language is not English. In fact, there are three such populations: a Puerto Rican group (along with some Blacks) in a large city in the Northeast, a Chicano group in a farming community in southern California, and a Cajun (French speaking) group in the South. Apparently when this Sponsor teaches the conceptual but only minimally verbal subject of math to these children whose native language is not English, there is evidence of strong gains. There is also reason to believe (based upon studies not reported herein) that the largest gains are made by the FT children in the non-verbal parts of the math test (the computation section of the test) rather than in the verbally-based math concepts. It is, of course, difficult to ascribe this finding to the lack of verbal content in the computations section of the test rather than to a curriculum which might be focused more on associative skills than on concept manipulation. Nevertheless, it is tempting to assert for future study that this Sponsor is able to teach math in a non-verbal mode to children who do not speak English as their first language. It would be important to know the success with which the Sponsor teaches math concepts to these children in their native language. The hypothesis that effective math instruction is, in this case, a function of the ease with which Spanish speaking children acquire math is not easily supported in these data because of the similar math effects found for very diverse cultural groups.

When we turn to the reading domain, the dependence of the effects on the kind of cultural group with which the Sponsor is involved is very apparent. Substantial reading gains are found in first grade for the two Spanish speaking groups, but not for the Cajun group. Clearly language instruction is aimed at the Spanish speaking, but math is not. Perhaps the efficacy of this Sponsor in teaching math has to do with the development of teacher sensitivity to the needs of children who find themselves in a school system dominated by a culture different from their own. The need to have the childrens' culture respected, to have the childrens' difficulty with English treated sensitively by those who can speak the childrens' native language, and have the childrens' adjustment to a different culture treated with understanding and support, might all be reasonably met by teachers trained in the bilingual mode. It is entirely possible that such teachers are able to be of very great support to any group of children learning math whose native language is not English regardless of what the native language is. This could be true even in those situations (e.g., Cajun) in which the language instruction is not fully worked out.

In any case, this Sponsor is showing strong effects in a variety of sites and with a variety of people. It is not clear why effects were not observed in kindergarten. It is possible that, as in the case of EDC, the kindergarten year could be an adaptation year. Given the very real problems these children face in adjusting to a strange school system, this is not an unreasonable hypothesis. Once again, preschool experience and future school effects should provide an opportunity to test the adaptation notion. In the meantime, it is clear that strong effects are emerging for SEDL that should be observed carefully in the future.

The conclusion we are suggesting here is that the pattern of effects which the first grade data reveal do not easily allow for hypotheses about differential effects of various curriculum models. Ther[illegible]e still no easily discernible patterns which correspond to any cluster of models. This is not to say that the Sponsors are not responsible for their effects. Rather, it is clear that each Sponsor interacts with the unique set of sites in a unique way. The particular kind of interaction present for each Sponsor is not yet available to us because there is too little data about the conditions at each site to allow for these generalizations. However, until such data do become available, it will be quite difficult to understand the variable patterns of effects and relate them to the intended inputs of each Sponsor.

This is not necessarily a pessimistic result for the policy maker who must make some decisions about providing guidance to local systems for improving the performance of their children. The goal of the Follow Through experiment was, from the beginning, to indicate the conditions under which each of the models generate each of its effects. The continued examination of each model at each site, with a large increase in the amount of site level data available, should provide this kind of guidance to local school systems, as well as a deeper understanding of the models themselves.

MONOGRAPH II

TEACHER STUDIES

MONOGRAPH II: TEACHER STUDIES

1.0 INTRODUCTION

The children in this study differ in a variety of ways: in family socioeconomic status, in parental participation, in treatment by Sponsors, and in the continuity of their presence in the data base. In order to describe more fully the dimensions of the Follow Through program impact, we have also examined the characteristics of the teachers who work with the children. It is essential that meaningful descriptors of the teachers be included. Teachers mediate the Sponsor's model; they tend to set the tone of the experiment as it impinges on children; and they contribute to the decisions which determine the formal shape of the program. We assume that teachers therefore influence pupil outcomes to varying degrees. Hence the teacher is one of the critical elements of the educational program which is delivered to each child.

Teachers differ with respect to two types of descriptors. The first type includes personal and professional characteristics of the teachers: age, ethnicity, years of experience in teaching, years of teaching in the same school, level of education, and type of certification. Many of these variables have played roles in other large studies of educational processes in the United States. Since these characteristics can be considered as generally external to the Sponsors' models, they are probably relatively independent of the intervention program. The second type of variables, on the other hand, relate more directly to the experimental manipulation: the training of the teachers in basic Sponsor philosophy; their attitudes toward specific philosophical and psychological principles; and their satisfaction with respect to the progress of various dimensions of parent involvement. This second class of variables can be considered internal to the sponsors' model and therefore as attributes of the delivery system.

This Monograph describes the teachers associated with the Follow Through program in terms of information collected with the Spring 1973 Teacher Questionnaire. Many of the tabular presentations include information pertaining to all teachers included in the survey: Cohort IV, kindergarten; Cohort III, non-entering first grade; Cohort III, second grade; Cohort II, second grade; and Cohort I, third grade. Since the Cohort III children, in the non-entering first grade, are the primary focus of this continuing study, we describe their teachers here. We present demographic information on the teachers working with other cohorts to indicate characteristics which are associated with each group of teachers. We cannot directly compare teachers representing the various cohorts: the cohort groups are too differently distributed across regions and city-size categories, so that differences between groups reflect more than merely cohort and grade distinctions (See Tables 1 and 2.)

The teacher "sample" is not a representative sample of any larger group. It is a group of teachers who, during the year 1972-1973, were teaching in FT or NFT classrooms and who completed the Spring 1973 Teacher Questionnaire. We cannot generalize from this group either to the complete population of FT teachers or to teachers in general. Since the data have been collected at a single point in time, furthermore, we have no way of determining whether variations (1) existed prior to the beginning of the experimental program, (2) resulted from a differential assignment of teachers to the treatment groups, or (3) represent the effects of the experimental manipulation. In describing the teachers' background characteristics and attitudes, we aim primarily to define those attributes which may mediate pupil performance. Even though we can generalize very little from our teacher data, an awareness of teacher characteristics is an important dimension of our understanding of pupil effects.

2.0 METHOD

2.1 The Sample

In the Spring of 1973, 1657 teachers representing 15 Sponsors' FT and NFT groups, responded to the Teacher Questionnaire mailed by Stanford Research Institute. The distribution of this total sample, shown in Table 3, indicates FT/NFT group membership by Sponsor, cohort, and grade level. From those surveyed, 1430 teachers (726 FT and 704 NFT) were selected according to the following criteria:

(1) each was associated with one of the ten Sponsors whose pupils were being studied; and

(2) each was the first teacher listed for a specific classroom.

TABLE MII-1

DISTRIBUTION OF TEACHERS IN FT/NFT STUDIES BY SPONSOR, COHORT, GRADE LEVEL AND REGION

	Northeast						North Central					
SPONSOR	Cohort IV, K		Cohort III, NEF		Other		Cohort IV, K		Cohort III, NEF		Other	
	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT
Far West	0	0	12	11	4	3	8	8	12	16	9	3
Arizona	11	4	19	9	7	2	16	16	22	26	2	2
Bank Street	12	9	25	42	28	21	0	0	0	0	0	0
Oregon	3	3	9	24	4	2	0	0	26	20	17	18
Kansas	18	14	24	25	30	6	9	4	15	14	3	3
High Scope	4	0	5	1	7	4	0	0	4	4	0	0
Florida	6	5	4	2	6	4	0	0	3	5	0	0
EDC	10	10	19	29	23	14	0	0	0	0	0	0
Pittsburgh	6	6	6	7	0	0	7	7	6	4	0	0
SEDL	9	8	9	9	15	0	0	0	0	0	0	0
Totals	79	59	132	159	124	56	40	35	88	89	31	26

	South						West					
SPONSOR	Cohort IV, K		Cohort III, NEF		Other		Cohort IV, K		Cohort III, NEF		Other	
	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT
Far West	0	0	0	0	0	0	14	13	24	18	9	1
Arizona	0	0	6	7	0	0	0	0	0	0	0	0
Bank Street	9	9	0	0	18	9	0	0	0	0	0	0
Oregon	0	0	0	0	16	7	0	0	0	0	0	0
Kansas	0	0	7	8	0	1	0	0	0	0	0	0
High Scope	4	14	4	10	10	11	3	3	15	16	0	1
Florida	0	0	14	20	24	64	0	0	8	3	0	0
EDC	2	0	2	7	11	17	0	0	0	0	0	0
Pittsburgh	0	0	0	0	0	0	0	0	0	0	0	0
SEDL	0	0	11	15	0	0	0	0	21	26	0	0
Totals	15	23	44	67	79	109	17	16	60	63	9	2

TABLE MII-2

DISTRIBUTION OF TEACHERS IN FT/NFT STUDIES BY SPONSOR, COHORT, GRADE LEVEL AND CITY SIZE

	< 10,000						≥ 10,000 < 50,000					
	Cohort IV, K		Cohort III, NEF		Other		Cohort IV, K		Cohort III, NEF		Other	
SPONSOR	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT
Far West	0	0	4	0	4	3	0	0	0	0	0	0
Arizona	0	0	0	2	0	0	11	4	22	9	4	1
Bank St	0	0	0	0	0	1	0	0	3	5	13	3
Oregon	0	0	2	3	7	3	0	0	0	0	9	4
Kansas	9	4	4	6	3	3	0	0	0	0	0	0
Hi Scope	0	1	0	3	0	2	7	16	7	8	10	9
Florida	0	0	3	5	0	0	0	0	8	3	6	8
EDC	0	0	3	7	10	17	0	0	1	5	2	1
Pittsburgh	7	7	6	4	0	0	6	6	6	7	0	0
SEDL	0	0	19	19	0	0	0	0	0	0	0	0
Totals	16	12	41	49	24	29	24	26	47	37	44	26

	≥ 50,000 < 200,000						≥ 200,000					
	Cohort IV, K		Cohort III, NEF		Other		Cohort IV, K		Cohort III, NEF		Other	
SPONSOR	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT
Far West	22	21	31	22	16	4	0	0	13	23	2	0
Arizona	16	16	11	13	0	0	0	0	14	18	5	3
Bank St	9	9	9	18	11	9	12	9	13	19	22	17
Oregon	0	0	28	37	17	18	3	3	5	4	4	2
Kansas	0	0	13	15	0	0	18	14	29	26	30	7
Hi Scope	0	0	0	0	0	0	4	0	21	20	7	5
Florida	0	0	0	0	14	18	6	5	18	22	10	42
EDC	0	0	5	11	0	0	12	10	12	13	22	13
Pittsburgh	0	0	0	0	0	0	0	0	0	0	0	0
SEDL	0	0	0	0	0	0	9	8	22	31	15	0
Totals	47	46	97	116	58	49	64	49	147	176	117	89

TABLE MII-3

DISTRIBUTION OF ORIGINAL TEACHER SAMPLE BY FT/NFT, SPONSOR, COHORT AND GRADE LEVEL

SPONSOR	Cohort IV Kindergarten		Cohort III Non-entering First		Cohort III Second Grade		Cohort IV Second Grade		Cohort II Third Grade		Cohort I Third Grade	
	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT	FT	NFT
Self-Sponsored	11	13	11	17	-	-	9	7	0	1	21	26
Far West	22	21	48	45	1	0	1	0	-	-	20	7
Arizona	27	20	47	42	-	-	-	-	0	1	9	3
Bank Street	21	18	25	42	5	4	10	4	4	3	27	19
Univ. of Georgia	-	-	-	-	6	0	1	0	-	-	-	-
Oregon	3	3	35	44	4	0	1	2	18	11	14	14
Kansas	27	18	46	47	-	-	17	4	0	1	16	5
High Scope	11	17	28	31	-	-	4	0	10	11	3	5
Florida	6	5	29	30	6	0	3	3	18	63	3	2
EDC	12	10	21	26	-	-	11	5	10	16	13	10
Pittsburgh	13	13	12	11	-	-	-	-	-	-	-	-
New York Univ.	4	5	2	4	-	-	5	2	-	-	4	3
SEDL	9	8	41	50	-	-	8	0	-	-	7	0
Southern Univ.	4	2	4	3	-	-	3	0	-	-	-	-
Calif. Dept. of Education	-	-	37	31	-	-	1	0	-	-	-	-
TOTALS	170	153	386	423	22	4	74	27	60	107	137	94

The distribution of these selected teachers is shown in Table 4.

Missing data on individual items reduced still further the teacher groups for which we can report distributions of background characteristics. Surviving samples ranged from N=1357 to N=1416, except for ethnicity, where N=945.

Complete data are available on all the program-related variables pertaining to teacher attitudes, training, satisfaction, and values for a subset of 498 FT and NFT teachers working in Cohort III, non-entering first grades. Table 5 shows the distribution of this selected group of teachers by FT/NFT and Sponsor.

2.2 Variable Development

The Teacher Questionnaire, administered in Spring 1973, contained approximately 240 items for the FT and NFT teachers to complete. A limited number of items, designed to focus on four research issues, were organized into the following categories for cluster analysis:

- attitudes of FT teachers toward Follow Through and Sponsor,
- focus of teacher training,
- reported teacher goals and practices, and
- parent involvement.

Data from a random sample of 400 teachers served as input to the cluster analysis procedure. First, a correlation matrix was computed for each of the four item categories using Kendall's tau. Then each of these matrices was submitted to cluster analysis using average linkage between merged groups. Finally, the method of Reciprocal Averages was employed to develop item weights for each item response in each of the resulting cluster variables which appear in Table 6.[1]

---

[1] For a more thorough presentation of the cluster analysis and scoring techniques, the reader is referred to our previous report, Cline et al. (1974: pp. I23-I35). Note, however, that there are several differences between this and last year's procedures: (1) correlations were computed on the whole sample, not averaged across subgroups, (2) no cross validation of cluster groups was performed, (3) Kendall's tau was used instead of Pearson r due to the assymetric distribution of many of the items, and (4) if any item in a cluster was missing, the entire cluster score was considered missing.

TABLE MII-4

DISTRIBUTION OF TEACHERS IN FT/NFT STUDIES BY SPONSOR, COHORT AND GRADE LEVEL

Sponsor	Cohort IV Kindergarten		Cohort III Non-entering First		Other*	
	FT	NFT	FT	NFT	FT	NFT
Far West	22	21	48	45	22	7
Arizona	27	20	47	42	9	4
Bank St	21	18	25	42	46	30
Oregon	3	3	35	44	37	27
Kansas	27	18	46	47	33	10
Hi Scope	11	17	28	31	17	16
Florida	6	5	29	30	30	68
EDC	12	10	21	36	34	31
Pittsburgh	13	13	12	11	-	-
SEDL	9	8	41	50	15	-
Totals	151	133	332	378	243	193

*Includes teachers in Cohort III, second grade; Cohort II, second and third grades; and Cohort I, third grade.

TABLE MII-5

DISTRIBUTION OF COHORT III NON-ENTERING FIRST GRADE TEACHERS IN STUDIES OF PROGRAM RELATED VARIABLES BY FT/NFT

SPONSOR	FT	NFT
Far West	39	31
Arizona	35	33
Bank Street	19	31
Oregon	25	31
Kansas	30	24
High Scope	21	19
Florida	26	19
EDC	13	23
Pittsburgh	10	5
SEDL	30	34
TOTALS	248	250

TABLE MII-6

SELECTED CLUSTER VARIABLES DEVELOPED FROM THE SPRING, 1973 TEACHER QUESTIONNAIRE

	Item #	Reliability*
FT--ONLY CLUSTER CENTER VARIABLES		
Teacher Satisfaction with FT/Sponsor		.95
Teacher would choose to teach in FT next year	11	
Teacher would recommend FT to a parent	12	
Teacher would advise friend to teach in FT	13	
Teacher would continue with Sponsor's approach	71	
Sponsor's approach suits teacher's style	72b	
Sponsor's approach meets pupils' learning needs	72d	
Amount of PAC activity		.90
Making decisions about school program	52a	
Making decisions about school budget	52b	
Helping select teachers	52c	
Designing parent workshops	52d	
Planning social activities	52e	
Mobilizing community support	52f	

FT/NFT CLUSTER VARIABLES

	Item #	Reliability
Amount of Teacher Training: Social/Emotional/ Intellectual Development and Individualization of Instruction		.93
Helping the child feel important	25a	
Developing enthusiasm for learning	25b	
Developing problem solving, and thinking critical	25d	
Developing child's ability to cooperate with peers	25e	
Helping the child make choices and become an independent learner	25f	
Diagnosing individual learning needs	25i	
Guiding children in individualized learning activities	25j	
Amount of Teacher Training: Classroom Management		.88
Cooperating effectively with other classroom adults	25k	
Working with small groups of children	25l	
Arranging the classroom environment	25m	
Maintaining discipline and control	25n	
Developing and/or selecting curriculum materials	25p	
Amount of Teacher Training: Parent Involvement		.90
Involving parents in children's learning	25g	
Involving parents in school operations	25h	
Reported Teacher Practices: Emphasis on Basic Skills vs. Affective Development		.64
Teacher places relatively little value on helping the child feel important	31a	

	Item #	Reliability
Teacher places relatively great value on developing basic skills	31c	
Teacher places emphasis on subject matter va. emotional needs	31d	
Reported Teacher Practices: Flexible / Open Classroom		.68
Children work independently vs. under adult supervision	32a	
Children choose their own activities and materials	32i	
Children interact freely	32k	
Children change places freely	32n	
Reported Teacher Practices: Individualization of Instruction		.51
Various activities occur simultaneously	32f	
Individual vs. group needs are dominant	32j	
Teacher seldom works with the whole class in reading, writing or math	33c	
Reported Teacher Practices: Teacher Centeredness		.60
Teacher imparts vs. child gathers information	32b	
Adults initiate interactions	32c	
Lesson plan is followed closely	32g	
Work and play are distinguished	32h	
Importance of Parent Involvement: Social/Educational Activities		.95
Meet the teacher	53a	
Participate in school's learning activities	53b	
Share ideas with teaching staff	53c	

	Item #	Reliability
Share cultural background with the children	53d	
Reinforce learning at home	53f	
Participate in parent-oriented workshops	53k	
Participate in school-oriented workshops	53l	
Plan and conduct social activities	53m	
Work with non-involved parents	53r	
Enlist support of community agencies	53o	
Importance of Parent Involvement: Decision Making Activities		[illegible]
Participate in decisions about school program	53g	
Participate in decisions about school budget	53i	
Help in selecting teacher assistants	53j	
Importance of Parent Involvement: Housekeeping Activities		.74
Help clean up	53d	
Help supervise the playground	53h	
Benefit of Parent Involvement		.83
How much parent involvement	54	
How helpful has parent involvement been to you	55a	
How helpful has parent involvement been to your children	55b	
How helpful has parent involvement been in bringing about community support for the program		
Teacher Satisfaction with Pupils' Affective Development		.79
Getting along with others	35c	
Understanding their heritage	35d	

	Item #	Reliability
Developing self-confidence	35e	
Developing an interest in school	35f	

---

*Coefficient alpha was used to estimate reliability coefficients.

In addition to the cluster variables, two items pertaining to teacher satisfaction were selected for further analyses. These included satisfaction with pupil progress in the areas of reading and reading readiness, as well as math and number concepts.

2.3 Analytic Techniques

Each of the cluster variables and single item variables which applied to both the FT and NFT teachers was submitted to a regression analysis using an analysis of covariance design in which treatment (FT/NFT) was nested within Sponsor. Teacher education, years of elementary school experience, and ethnicity served as covariates. Table 7 presents the complete coding scheme employed; it is designed to allow us to test the impact of each Sponsor's program on teacher attitudes and behaviors with teacher background variables held constant.

3.0 RESULTS

3.1 Personal and Professional Characteristics

We have chosen to report teacher's education and experience, salary, age, and ethnicity primarily because these measures have theoretical implications for both teacher and pupil outcomes. The reader is referred to our previous report, Cline, (1974: Vol. I, pp. II-3 to II-5) for a complete discussion of this rationale.

3.1.1 Age and Experience

Tables 8 through 10 indicate the mean age, the total length of elementary school teaching experience, and the length of teaching experience in the current school for the FT and the NFT teachers, by Sponsor, and cohort/grade. Table 11 displays the mean number of years the FT teachers associated with each Sponsor have spent using their Sponsor's approach, by Sponsor and cohort/grade.

In general, these teachers are in their mid-thirties, have over nine years total teaching experience, and have been teaching over five years in their current schools.[2] There are some variations by Sponsor and by cohort/

---

[2] When we compare the age, experience, and education of the teachers in our sample with two nationally representative samples, we find that they are fairly comparable. One sample is the set of teachers in elementary schools attended by the average Black and White pupil (Coleman et al., 1966). The

TABLE MII-7
PREDICTOR CODING SCHEME FOR THE NESTED ANALYSIS OF TEACHER OUTCOMES

SPONSOR	FT/NFT Status	10 FT-NFT Effects Within Sponsors										9 Contrasts Among Sets of Sponsors								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Far West	FT	.5	0	0	0	0	0	0	0	0	0	1	-4	4	0	0	0	0	0	0
	NFT	-.5	0	0	0	0	0	0	0	0	0	1	-4	4	0	0	0	0	0	0
Arizona	FT	0	.5	0	0	0	0	0	0	0	0	1	1	-1	1	1	1	0	0	0
	NFT	0	-.5	0	0	0	0	0	0	0	0	1	1	-1	1	1	1	0	0	0
Bank St	FT	0	0	.5	0	0	0	0	0	0	0	1	1	-1	1	1	-1	0	0	0
	NFT	0	0	-.5	0	0	0	0	0	0	0	1	1	-1	1	1	-1	0	0	0
Oregon	FT	0	0	0	.5	0	0	0	0	0	0	1	1	-1	-1	-1	0	1	0	0
	NFT	0	0	0	-.5	0	0	0	0	0	0	1	1	-1	-1	-1	0	1	0	0
Kansas	FT	0	0	0	0	.5	0	0	0	0	0	1	1	-1	-1	-1	0	-1	0	0
	NFT	0	0	0	0	-.5	0	0	0	0	0	1	1	-1	-1	-1	0	-1	0	0
High Scope	FT	0	0	0	0	0	.5	0	0	0	0	-1	-4	-4	0	0	0	0	0	0
	NFT	0	0	0	0	0	-.5	0	0	0	0	-1	-4	-4	0	0	0	0	0	0
Florida	FT	0	0	0	0	0	0	.5	0	0	0	-1	1	1	1	-1	0	0	1	0
	NFT	0	0	0	0	0	0	-.5	0	0	0	-1	1	1	1	-1	0	0	1	0
EDC	FT	0	0	0	0	0	0	0	.5	0	0	-1	1	1	1	-1	0	0	-1	0
	NFT	0	0	0	0	0	0	0	-.5	0	0	-1	1	1	1	-1	0	0	-1	0
Pittsburgh	FT	0	0	0	0	0	0	0	0	.5	0	-1	1	1	-1	1	0	0	0	1
	NFT	0	0	0	0	0	0	0	0	-.5	0	-1	1	1	-1	1	0	0	0	1
SEDL	FT	0	0	0	0	0	0	0	0	0	.5	-1	1	1	-1	1	0	0	0	-1
	NFT	0	0	0	0	0	0	0	0	0	-.5	-1	1	1	-1	1	0	0	0	-1

TABLE MII-8
MEANS AND STANDARD DEVIATIONS FOR
TEACHER AGE
BY FT/NFT, SPONSOR AND COHORT/GRADE

SPONSOR	Kindergarten, Cohort IV						Non-entering First, Cohort III						Other*					
	FT			NFT			FT			NFT			FT			NFT		
	N	X	SD	N	X	SD	N	X	SD	N	X	SD	N	X	SD	N	X	SD
Far West	22	37.27	13.40	21	44.29	12.71	47	38.13	12.76	45	43.93	13.65	21	34.62	9.54	7	33.29	10.61
Arizona	26	40.73	14.70	20	37.00	9.42	47	33.00	9.86	40	38.20	14.06	9	37.89	12.59	4	32.50	4.36
Bank St	17	33.76	14.97	16	41.06	14.49	23	34.78	13.83	39	37.28	14.29	45	35.16	10.90	29	38.59	12.41
Oregon	3	40.67	15.95	2	37.00	18.38	34	29.65	7.06	43	38.35	11.89	36	32.11	11.27	26	32.96	11.01
Kansas	23	30.52	8.71	18	33.94	10.69	44	32.64	10.81	43	40.05	14.15	27	32.00	12.33	10	33.30	9.32
Hi Scope	10	33.60	10.07	17	44.65	9.56	28	38.29	11.84	30	39.20	12.43	16	32.94	10.85	16	33.75	10.13
Florida	6	30.00	6.78	5	36.20	13.16	25	31.92	8.95	28	38.93	13.22	28	34.04	12.00	68	39.69	12.39
EDC	10	36.30	10.72	9	44.89	12.32	20	31.85	10.33	29	37.93	13.52	31	35.13	11.79	29	41.10	12.46
Pittsbrgh	13	40.54	12.81	13	36.31	10.75	12	34.00	11.53	11	43.91	14.64	--	--	--	--	--	--
SEDL	7	28.71	9.76	8	26.38	4.34	41	35.71	11.69	48	34.02	10.52	12	37.42	14.22	--	--	--
Totals	137	35.65	12.72	129	39.06	12.12	321	34.16	11.17	356	38.81	13.26	225	34.17	11.45	189	37.58	11.97

*Includes teachers in Cohort III, second grade; Cohort II, second and third grades; and Cohort I, third grade.

TABLE MII-9

MEANS AND STANDARD DEVIATIONS FOR YEARS OF TOTAL TEACHING EXPERIENCE BY FT/NFT, SPONSOR AND COHORT/GRADE

SPONSOR	Kindergarten, Cohort IV						Non-entering First, Cohort III						Other*					
	FT			NFT			FT			NFT			FT			NFT		
	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD
Far West	22	10.46	12.09	21	17.52	14.27	48	10.40	7.86	45	16.54	12.29	22	8.27	8.30	7	6.57	6.08
Arizona	27	14.93	16.88	20	10.55	6.27	47	7.00	7.60	42	12.71	11.61	9	7.00	5.64	4	9.00	3.83
Bank St	21	10.05	10.13	18	17.00	12.66	25	9.00	8.61	42	10.95	9.92	45	9.87	8.41	29	13.17	9.48
Oregon	3	7.67	3.22	3	11.00	7.55	33	5.94	4.37	43	11.05	9.37	36	7.58	7.26	27	8.07	7.74
Kansas	27	6.37	5.52	18	6.94	6.14	46	8.46	7.97	46	12.76	9.38	33	10.12	10.54	10	10.00	8.03
High Scope	11	7.82	5.76	17	16.94	8.22	28	8.86	5.56	31	11.77	8.37	16	8.06	8.82	16	7.75	6.43
Florida	6	4.17	4.31	5	11.80	11.28	29	8.69	8.26	28	13.86	11.27	30	8.33	8.85	68	12.62	10.31
EDC	12	8.42	5.57	10	13.90	8.65	20	7.75	8.46	34	11.71	9.20	34	9.50	9.19	31	15.61	10.32
Pittsburgh	13	20.31	22.82	13	11.08	8.87	12	10.83	10.90	11	13.91	9.19	-	-	-	-	-	-
SEDL	9	3.78	2.49	8	4.38	1.92	41	8.83	8.73	50	8.46	7.71	14	7.57	5.36	-	-	-
Totals	151	10.26	12.50	133	12.84	10.29	329	8.46	7.78	372	12.18	10.09	239	8.80	8.47	192	11.71	9.52

* Includes teachers in Cohort II, second grade; Cohort II, second and third grades; and Cohort I, third grade.

TABLE MII-10

MEANS AND STANDARD DEVIATIONS FOR YEARS OF TEACHING EXPERIENCE
IN CURRENT SCHOOL BY FT/NFT, SPONSOR AND COHORT/GRADE

SPONSOR	Kindergarten, Cohort IV						Non-entering First, Cohort III						Other*					
	FT			NFT			FT			NFT			FT			NFT		
	N	X̄	SD	N	X̄	SD	N	X̄	SD	N	X̄	SD	N	X̄	SD	N	X̄	SD
Far West	22	5.14	4.61	21	6.00	6.99	47	6.34	5.52	45	10.07	8.91	22	5.54	4.98	7	3.14	2.41
Arizona	27	6.78	8.56	20	4.65	2.87	47	3.83	4.38	42	7.74	8.18	9	6.11	5.73	4	5.00	3.46
Bank St	21	3.86	2.85	18	6.72	4.46	25	6.56	7.55	42	6.31	6.33	45	6.61	7.11	29	8.31	6.62
Oregon	3	3.00	2.65	3	7.67	4.04	35	3.60	3.08	44	6.66	6.04	36	5.00	6.12	27	5.07	4.57
Kansas	26	5.46	6.26	18	5.56	5.23	45	4.91	4.31	46	8.11	7.19	31	5.48	4.80	10	9.20	7.02
High Scope	11	4.73	4.47	17	7.65	7.90	28	4.96	4.29	31	6.81	5.11	17	4.76	4.41	16	4.00	3.46
Florida	6	1.67	1.86	5	7.60	5.60	29	5.00	4.82	28	6.00	7.31	29	4.28	5.96	68	6.40	5.92
EDC	12	3.67	2.39	10	9.90	7.36	20	2.85	3.31	34	8.76	8.55	34	5.71	7.04	31	12.52	10.76
Pittsburgh	13	9.31	8.09	13	4.69	3.52	12	7.42	8.98	11	9.00	6.54	-	-	-	-	-	-
SEDL	9	3.00	1.73	8	2.75	2.12	41	6.34	6.74	50	5.98	6.10	15	5.46	4.29	-	-	-
Totals	150	5.21	5.84	133	6.11	5.60	329	5.10	5.34	373	7.46	7.21	238	5.49	5.93	192	7.29	7.13

*Includes teachers in Cohort II, second grade; Cohort II, second and third grades; and Cohort I, third grade.

TABLE MII-11

MEANS AND STANDARD DEVIATIONS FOR NUMBER OF YEARS IN SPONSOR'S APPROACH BY SPONSOR AND COHORT/GRADE

Sponsor	Kindergarten, Cohort IV			Non-entering 1st, Cohort III			Other*		
	N	X	SD	N	X	SD	N	X	SD
Far West	22	3.50	1.57	48	2.92	1.30	19	2.21	.98
Arizona	25	2.52	1.45	45	2.76	3.51	7	3.00	1.56
Bank St	21	3.19	1.60	24	2.42	1.53	44	2.43	1.17
Oregon	3	1.00	0	34	2.65	1.28	37	2.35	1.23
Kansas	26	3.04	1.15	45	2.64	1.13	33	2.15	1.00
Hi Scope	10	3.50	2.88	28	2.46	1.10	17	2.94	.25
Florida	6	1.83	1.72	29	2.66	1.14	30	1.77	1.28
EDC	12	2.92	1.73	20	1.85	1.42	33	1.88	1.05
Pittsbrgh	13	3.00	1.22	12	2.42	1.56	--	--	--
SEDL	9	2.44	1.81	40	3.28	6.08	15	2.60	1.24
Totals	147	2.93	1.63	325	2.69	2.72	235	2.26	1.15

*Includes teachers in Cohort III, second grade; Cohort II, second and third grades; and Cohort I, third grade.

grade across these variables. Further, while the FT teachers tend to be somewhat younger, have fewer total years of experience, and have been teaching fewer years in their current schools, they are not inexperienced teachers. In the case of the Cohort III teachers, the FT/NFT contrasts remain fairly consistent across Sponsors, with two exceptions. SEDL FT teachers are slightly older and have more experience both in teaching and in their current schools. Bank Street also tends to have FT teachers with more years of experience in their current schools.

While Sponsors vary somewhat in their teachers' level of experience with their approach, the Cohort III teachers, on the average, have slightly over two and one-half years of experience. EDC has the least experienced group, with an average of 1.85 years of experience with that approach; SEDL has the most experienced group, with an average of 3.28 years of experience.

3.1.2 Education and Certification

Table 12 indicates the percentages of teachers by Sponsor, FT/NFT, and cohort/grade with a Bachelor's degree or less, with more than a Bachelor's degree but less than a Master's degree, and with a Master's degree or more. Table 13 shows the percentages of teachers with temporary, renewable, or permanent certification.

FT and NFT teachers differ very little in educational attainment. The cohort/grade groups show more variability: approximately 50 per cent of the Cohort II, entering-first teachers have earned a Bachelor's degree or more, in contrast to 65 per cent of the teachers in the other cohort/grade groups. The Sponsor groups also vary considerably: among the Cohort III teachers, the NFT teachers tend to have more education than the FT teachers. In Kansas and High/Scope, however, more FT teachers than NFT have earned more than a Bachelor's degree. In the other cohort/grade groups, FT teachers tend to have a higher educational attainment level, but the Sponsors vary substantially.

The FT and NFT teacher groups differ in their levels of certification. For example, 71.4 per cent of the Cohort III NFT teachers are permanently

---

other is the sample of elementary teachers in low SES schools studied by Herrioh and St. John (1966).

TABLE MII-12

PERCENTAGE OF TEACHERS WITH BACHELOR'S DEGREES OR LESS, MORE THAN A BACHELOR'S DEGREE BUT LESS THAN A MASTER'S DEGREE, OR MASTER'S DEGREE OR MORE BY FT/NFT, SPONSOR, AND COHORT/GRADE

SPONSORS		Kindergarten, Cohort IV				Non-entering First, Cohort III				Other*			
			3	4	5		3	4	5		3	4	5
		N	%	%	%	N	%	%	%	N	%	%	%
Far West	FT	22	36.4	50.0	13.6	48	56.3	37.5	6.3	22	36.4	50.0	13.6
	NFT	21	38.1	47.6	14.3	44	52.3	34.1	13.6	7	57.1	28.6	14.3
Arizona	FT	27	40.1	44.4	7.4	47	36.2	42.6	21.3	9	55.6	33.3	11.1
	NFT	20	45.0	35.0	20.0	42	35.7	35.7	28.6	4	-	75.0	25.0
Bank St	FT	20	30.0	40.0	30.0	25	44.0	24.0	32.0	45	26.7	46.7	26.7
	NFT	18	27.8	61.1	11.1	41	34.1	41.5	24.4	28	28.6	35.7	35.7
Oregon	FT	3	-	33.3	66.7	25	34.3	48.6	17.1	36	19.4	61.1	19.4
	NFT	3	-	-	00.0	44	22.7	40.9	36.4	27	37.0	40.7	22.2
Kansas	FT	27	29.6	44.4	25.9	47	32.6	37.0	30.4	33	24.2	45.5	30.3
	NFT	18	44.4	27.8	27.8	45	44.4	35.6	20.0	10	10.0	40.0	50.0
Hi Scope	FT	11	27.3	36.4	36.4	28	40.7	51.9	7.4	17	23.5	47.1	29.4
	NFT	17	56.8	29.4	11.8	30	53.3	30.0	16.7	16	56.3	37.5	6.3
Florida	FT	6	33.3	66.7	-	29	44.8	37.9	17.2	30	53.3	33.3	13.3
	NFT	5	40.0	20.0	40.0	28	32.1	39.3	28.6	65	58.5	27.7	13.8
EDC	FT	10	20.0	20.0	60.0	20	40.0	30.0	30.0	32	45.5	33.3	18.2
	NFT	10	20.0	30.0	50.0	34	35.3	44.1	20.6	30	63.3	23.3	13.3
Pittsbrgh	FT	13	61.5	23.1	15.4	12	91.7	8.3	-	-	-	-	-
	NFT	13	53.8	46.2	-	11	90.9	9.1	-	-	-	-	-
SEDL	FT	9	22.2	33.3	44.4	41	60.9	24.4	14.6	15	20.0	46.7	33.3
	NFT	8	-	62.5	37.5	49	57.1	32.7	10.2	-	-	-	-
TOTALS	FT	148	35.1	40.6	24.3	330	48.2	34.2	17.6	240	34.2	44.1	21.7
	NFT	136	45.8	35.0	31.2	370	45.8	34.3	19.9	787	38.8	38.5	22.7

KEY

3 = Bachelor's Degree or less
4 = More than a Bachelor's Degree
5 = Master's Degree or more

*Includes teachers in Cohort III, second grade; Cohort II, second and third grades; and Cohort I, third grade.

TABLE MII-13

PERCENTAGE OF TEACHERS WITH TEMPORARY, RENEWABLE, OR PERMANENT CERTIFICATION BY FT/NFT, SPONSOR, AND COHORT/GRADE

SPONSORS		Kindergarten, Cohort IV				Non-entering First, Cohort III				Other **			
		N	2 %	3 %	4 %	N	2 %	3 %	4 %	N	2 %	3 %	4 %
Far West	FT	22	13.6	40.9	45.5	45	13.3	24.4	62.2	22	9.1	50.0	40.9
	NFT	21	9.5	47.6	42.9	42	2.4	21.4	76.2	7	-	42.9	57.1
Arizona	FT	26	7.7	15.4	76.9	47	6.5	38.3	53.2	9	22.2	11.1	66.7
	NFT	20	-	30.0	70.0	42	2.4	33.3	64.3	4	25.0	50.0	25.0
Bank St	FT	20	20.0	30.0	50.0	23*	17.4	4.2	78.3	44	2.3	15.9	81.8
	NFT	18	16.7	5.6	77.8	40*	20.0	7.5	72.5	30	6.7	10.0	83.3
Oregon	FT	3	-	-	100.0	32*	34.4	12.5	53.1	37	16.2	24.3	59.5
	NFT	3	-	-	100.0	38*	21.0	13.2	65.8	24	25.0	33.3	41.7
Kansas	FT	25	16.0	8.0	76.0	45	11.1	13.3	75.6	30	23.3	3.3	73.3
	NFT	18	38.9	-	61.1	43	7.0	16.3	76.7	10	10.0	-	90.0
Hi Scope	FT	11	9.1	27.3	63.6	27	7.4	40.7	51.9	17	11.8	52.9	35.3
	NFT	17	-	70.6	29.4	30	-	43.3	56.7	15	-	60.0	40.0
Florida	FT	6	33.3	-	66.7	29	34.5	13.8	51.7	30	3.3	46.7	50.0
	NFT	5	40.0	-	60.0	25	16.0	16.0	68.0	65	7.7	49.2	43.1
EDC	FT	12	16.7	-	83.3	20	20.0	5.0	75.0	31	16.1	25.8	58.1
	NFT	10	-	-	100.0	34	14.7	8.8	76.5	31	6.5	25.8	67.7
Pittsburgh	FT	13	30.8	-	69.2	12	16.7	41.7	41.7	-	-	-	-
	NFT	13	15.4	38.5	46.2	11	-	18.2	81.8	-	-	-	-
SEDL	FT	9	33.3	11.1	55.6	36	16.7	19.4	63.9	14	21.4	21.4	57.1
	NFT	8	25.0	-	75.0	50	8.0	16.0	76.0	-	-	-	-
TOTALS	FT	147	18.0	13.3	68.7	313	18.0	21.3	60.7	234	14.0	27.9	58.1
	NFT	133	14.6	19.2	66.2	355	9.2	19.4	71.4	186	10.1	33.9	56.0

KEY
2 = temporary
3 = renewable
4 = permanent

* Cells include one uncertified teacher each

** Includes teachers in Cohort II, second grade; Cohort II, second and third grades; and Cohort I, third grade.

certified, as opposed to 60.7 per cent of the FT teachers. In addition, the average percentages of FT and NFT teachers with renewable certification are 21.3 per cent and 19.4 per cent, respectively; and the average percentages of FT and NFT teachers with temporary certification are 18.0 per cent and 9.2 per cent respectively. There is considerable variability among Sponsors on these certification patterns, however.

Bank Street and EDC have FT/NFT groups with over 70 per cent permanently certified teachers and approximately 20 per cent holding temporary certification. Fewer than 8 per cent have renewable certification. A different pattern exists for Oregon and Florida. Both Sponsors are using approximately 50 per cent non-permanently certified FT teachers, two-thirds of whom have temporary certification, only. By comparison, approximately 65 per cent of the NFT teachers are permanently certified. In Arizona, High/Scope, and Pittsburgh, FT teachers tend to have either permanent or renewable certification. More NFT comparison teachers associated with these sponsors hold permanent certification. The distributions by region and city size of Sponsors with parallel certification patterns indicates that these patterns may be community related. (It also suggests that the FT/NFT matching did a good job of eliminating the potential confounding effects of certification level, if any.) For example, the proportions of FT teachers in Bank Street and EDC are similarly distributed, with the greatest proportion in large cities in the Northeast.

In general, the FT/NFT sample distributions on certification indicate that those teachers who are not permanently certified, i.e. "marginal teachers" are either assigned or attracted to the FT experiment. Even though the FT teachers tend to be experienced, and have attained a fairly high level of education, they tend to be somewhat marginal with respect to the teaching profession. This varies by Sponsor, but it does reveal FT's generally perceived experimental status.

3.1.3. Salary

Table 14 indicates the mean salary level of the teachers by FT/NFT, Sponsor, and cohort/grade. The NFT teachers tend to earn slightly more than the FT teachers. This is possibly a reflection of the different experience and certification levels of the two groups. This pattern is con-

TABLE MII-14

MEANS AND STANDARD DEVIATIONS FOR
TEACHER SALARY
BY FT/NFT, SPONSOR AND COHORT/GRADE

SPONSOR	Kindergarten, Cohort IV						Non-entering First, Cohort III						Other *					
	FT			NFT			FT			NFT			FT			NFT		
	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD
Far West	22	5.50	1.99	20	5.50	2.86	47	6.36	2.62	44	7.57	2.88	22	6.14	2.40	7	5.57	2.44
Arizona	27	6.22	2.89	20	5.60	2.06	47	5.62	2.35	41	5.83	2.19	9	6.78	1.99	4	5.50	1.00
Bank St	19	2.90	1.70	16	8.31	1.82	21	7.00	2.95	41	7.39	2.31	44	6.91	2.79	28	8.86	2.44
Oregon	3	8.67	3.06	2	11.00	2.83	33	6.06	2.40	43	7.63	2.40	36	4.97	2.13	26	5.92	2.50
Kansas	26	6.69	2.65	18	7.06	3.96	43	6.21	2.27	44	6.75	2.61	25	7.20	2.26	10	6.60	2.17
Hi Scope	11	6.73	3.55	17	5.65	1.22	28	6.54	2.20	31	6.68	2.60	16	4.75	3.09	16	5.00	3.44
Florida	5	6.80	3.27	5	8.20	4.87	28	5.75	1.80	28	5.93	1.98	28	4.64	2.04	68	5.28	1.70
EDC	10	7.70	2.11	10	9.30	1.70	20	6.50	2.82	33	6.76	3.03	31	5.90	2.14	30	7.50	2.66
Pittsbrgh	13	6.15	2.41	13	5.15	1.95	12	5.83	1.80	11	6.18	1.40	--	--	--	--	--	--
SEDL	8	6.25	3.15	8	6.00	1.51	39	5.87	2.60	48	6.15	2.37	14	7.79	2.39	--	--	--
Totals	144	6.63	2.63	129	6.58	2.84	318	6.13	2.42	364	6.76	2.54	225	6.03	2.56	189	6.31	2.64

KEY

1 = Below $5000
2 = $5000 - 5999
3 = 6000 - 6999
4 = 7000 - 7999
5 = 8000 - 8999
6 = 9000 - 9999
7 = $10,000 - 10,999
8 = 11,000 - 11,999
9 = 12,000 - 12,999
10 = 13,000 - 13,999
11 = 14,000 - 14,999

*Includes teachers in Cohort II, second grade; Cohort II, second and third grades; and Cohort I, third grade.

sistent across teachers for all Sponsors in the Cohort III group. The Cohort IV kindergarten teachers tend to be more highly paid, probably because they teach in the big cities, where salaries tend to be higher. These Cohort IV teachers also tend to be more experienced and have a higher degree of educational attainment.

3.1.4 Ethnicity

Table 15 indicates the percentage of White teachers in the FT/NFT sample by Sponsor and cohort/grade. Overall, there are more minority teachers in the FT than in the NFT classes. In Cohort III, teacher ethnicity varies considerably among the Sponsors, paralleling, in some cases, that of the pupils. For example, Pittsburgh FT/NFT pupils and teachers are both over 98 per cent White, while SEDL pupils and teachers are both slightly less than 50 per cent White. In other instances, such as Oregon with 5.3 per cent White FT pupils vs. 71.4 per cent White teachers and 19.5 per cent White NFT pupils vs. 79.5 per cent White teachers, there is a considerable difference. Unfortunately, without merging teacher/pupil data, we do not know in what way teacher-pupil ethnic match, or any of the other teacher background variables, mediate Sponsor impact on pupil performance.

3.2 Teacher Satisfaction with FT Sponsor

Table 16 displays the means and standard deviations for the FT teachers on this cluster variable. It should be noted that this cluster is constructed from items which solicit the teachers' opinions about the suitability of the model to their pupils, themselves, and other teachers. This cluster variable, therefore, contains global judgements about the usefulness of the Sponsor's program. In general, no Sponsor group differs greatly from the overall FT mean. There appears to be a general degree of teacher satisfaction with each of the models. There are, however, some differences among Sponsor groups which are worthy of note.

The teachers in Cohort III who are associated with Bank Street and Pittsburgh express the highest degree of satisfaction with their Sponsor's program. The teachers associated with Oregon express the least degree of satisfaction. It is interesting that teachers in two programs which emphasize a highly individualized application of teacher skill and training (Bank Street), and curriculum materials (Pittsburgh), indicate more satis-

TABLE MII-15

PERCENTAGE OF WHITE TEACHERS BY FT/NFT, SPONSOR, AND COHORT/GRADE

SPONSORS	Kindergarten, Cohort IV			Non-entering First, Cohort III			Other*		
	N	FT %	NFT %	N	FT %	NFT %	N	FT %	NFT %
Far West	42	100.0	100.0	61	70.8	60.0	21	71.4	85.7
Arizona	44	88.9	100.0	77	80.9	92.9	10	77.8	75.0
Bank Street	17	58.8	38.9	49	90.9	69.0	31	39.5	46.7
Oregon	3	33.3	100.0	60	71.4	79.5	40	62.2	63.0
Kansas	28	66.7	55.6	55	51.1	72.7	23	48.4	80.0
Hi Scope	25	72.7	100.0	43	70.4	77.4	16	46.7	56.3
Florida	4	60.0	20.0	27	55.2	40.7	68	71.4	72.7
EDC	13	70.0	60.0	32	50.0	64.7	41	68.8	61.3
Pittsburgh	26	100.0	100.0	23	100.0	100.0	-	-	-
SEDL	11	37.5	100.0	48	48.8	56.0	7	50.0	-
TOTALS	213	69.8	77.4	475	69.0	71.3	257	59.6	67.6

* Includes teachers in Cohort III, second grade; Cohort II, second and third grades; and Cohort I, third grade.

TABLE MII-16

MEANS AND STANDARD DEVIATIONS FOR TEACHERS' SATISFACTION BY SPONSOR AND COHORT/GRADE

SPONSOR	Kindergarten, Cohort IV			Non-entering First, Cohort III			Other*		
	N	$\overline{X}$	SD	N	$\overline{X}$	SD	N	$\overline{X}$	SD
Far West	22	37.227	4.566	48	32.667	8.667	21	33.286	6.076
Arizona	25	32.000	9.078	44	32.750	7.308	7	31.714	8.616
Bank St	19	35.053	4.927	24	37.458	4.374	42	32.500	6.685
Oregon	2	32.000	12.728	31	31.226	8.249	35	32.400	6.103
Kansas	26	35.885	4.246	41	32.244	7.625	26	32.346	7.310
Hi Scope	11	34.818	5.845	27	35.185	7.211	15	32.533	8.123
Florida	5	36.000	9.028	29	36.448	4.702	28	30.214	9.670
EDC	10	35.400	5.275	18	33.778	5.375	32	32.656	7.334
Pittsbrgh	13	30.538	14.892	12	37.500	4.758	--	------	-----
SEDL	9	33.667	10.210	38	32.895	5.802	13	31.154	7.105
Totals	142	34.500	7.763	312	33.696	7.122	219	32.169	7.265

*Includes teachers in Cohort III, second grade; Cohort II, second and third grades; and Cohort I, third grade.

faction than those teachers (Oregon) who rely primarily on a structured curriculum designed for use with groups of children. Whether the teachers are expressing more satisfaction about the model suiting their style or suiting the needs of the pupils is not clear, since this variable is a combination of these and other facets associated with the model.

The expressed satisfaction of the Cohort III teachers is not entirely consistent with the Cohort IV group or with the "other" group. In the latter two groups, Far West teachers express more satisfaction than other Sponsors' teachers and Pittsburgh Cohort IV kindergarten teachers express the least satisfaction. Those who are least satisfied in the "other" group are associated with the Florida model. It is not possible to identify the basic reasons for the differences in expressed satisfaction. As was mentioned previously, these groups are not comparable. The distributions of teachers by region and city size are very different. The differences in the displayed results could be attributed to either demographic or grade level factors. For example, the teachers of Cohort IV kindergarten children could be expressing satisfaction/dissatisfaction with their Sponsors' programs when they are applied to a group of pupils in the big cities. On the other hand, they may be more or less satisfied with kindergarten components of the Sponsors' programs which are substantively different from the components being used by teachers in the other grades.

3.3 Amount of Teacher Training

Three cluster variables were developed to investigate the amount of teacher training reported in:

- Social/emotional/intellectual development and individualization of instruction,
- Classroom management, and
- Parent involvement.

Tables 17 through 19 present data which indicate the contrasts between the FT and the NFT teachers. In general, an arrow pointing upward indicates that the FT group of teachers exceeds the NFT group of teachers on that variable. An arrow pointing downward indicates the reverse is true. All differences are reported in standard deviation units of the outcome variable for the purpose of comparability.

Table 17 indicates that the teachers associated with Far West and Bank Street report receiving significantly more training in social/emotional/intellectual development and individualization of instruction than their respective NFT teachers. Only Florida and Pittsburgh FT teachers report receiving less training on this dimension than their NFT comparisons. Their NFT groups are receiving more training than the other NFT groups on this dimension.

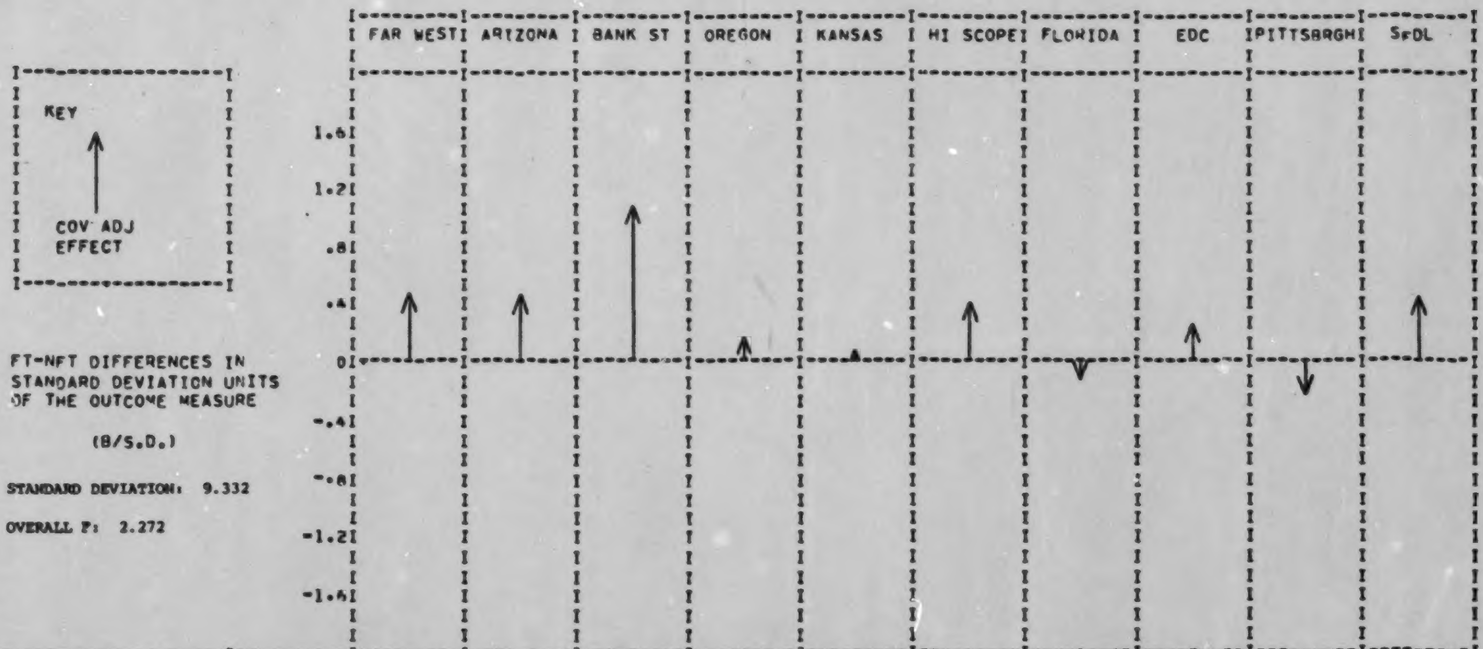
Table 18 displays the contrasts between the FT and NFT teachers on amount of training reported in classroom management. Arizona, Bank Street, and Kansas place considerably more emphasis on this dimension of training. The FT teachers in Florida and EDC report receiving somewhat less training in classroom management than their NFT comparisons.

Table 19 shows the comparisons in the amount of training in parent involvement which the FT and the NFT teachers report receiving. The teachers associated with Bank Street, Kansas, Florida, EDC, and SEDL indicate they are receiving significantly more training in parent involvement than the NFT's for each Sponsor. It is noteworthy that no Sponsor's teachers report receiving less training in parent involvement than the NFT's.

Table 20 shows the comparisons in the amount of teacher training on all three cluster variables. It is clear from the profile that the Sponsors differ in both the amount of training and the content areas which receive emphasis. FT teachers in all Sponsor programs report receiving more training than the NFT teachers. The patterns reported by the teachers, however, fluctuate. In some cases the variations reflect Sponsor orientations. For example, Oregon, which places considerably more emphasis on using prescribed curriculum materials than on teacher training, has teachers reporting that they receive relatively little training. Florida has a heavy emphasis on parent involvement in the curriculum delivery; this is reflected in the teachers' reports in the relative amount of training placed on this component. In some cases they do not. The Pittsburgh FT teachers could be expected to show a positive contrast with the NFT teachers on the social/emotional/intellectual development and individualization of instruction component. Apparently, the items pertaining to individualization of instruction are overshadowed by the other variables entering the cluster.

TABLE MII-17

SPONSOR PROFILE FOR AMOUNT OF TEACHER TRAINING: SOCIAL/EMOTIONAL/INTELLECTUAL DEVELOPMENT AND INDIVIDUALIZATION OF INSTRUCTION



STANDARD DEVIATION: 9.332

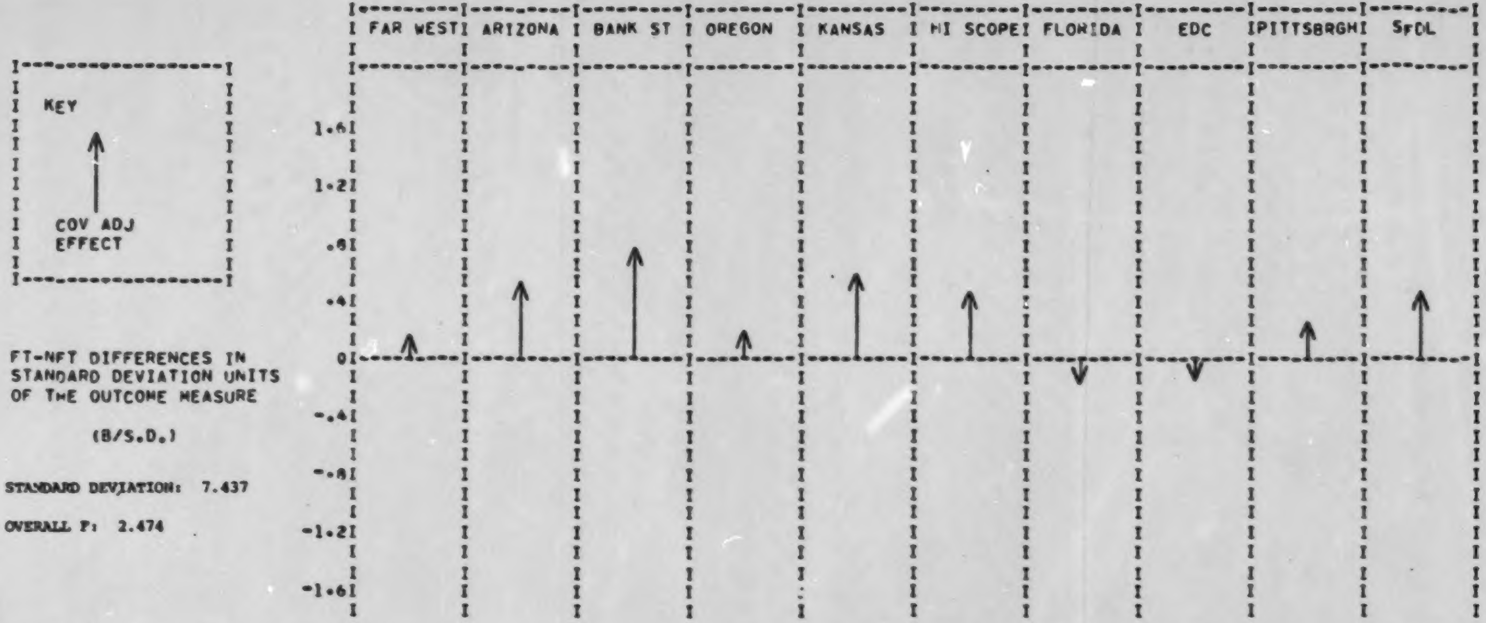
OVERALL F: 2.272

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		4.305	4.104	10.091	1.273	.095	3.375	-.904	1.921	-2.224	4.210
S.E. = STANDARD ERROR OF B		2.182	2.215	2.627	2.433	2.494	2.854	2.728	3.132	4.928	2.258
F= SIGNIFICANCE STATISTIC		* 3.893	3.434	* 14.759	.274	.001	1.399	.130	.376	.204	3.478
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		.461	.440	1.075	.126	.010	.362	-.105	.206	-.238	.451
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

* p less than .05

TABLE MII-18

SPONSOR PROFILE FOR AMOUNT OF TEACHER TRAINING: CLASSROOM MANAGEMENT



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

(B/S.D.)

STANDARD DEVIATION: 7.437

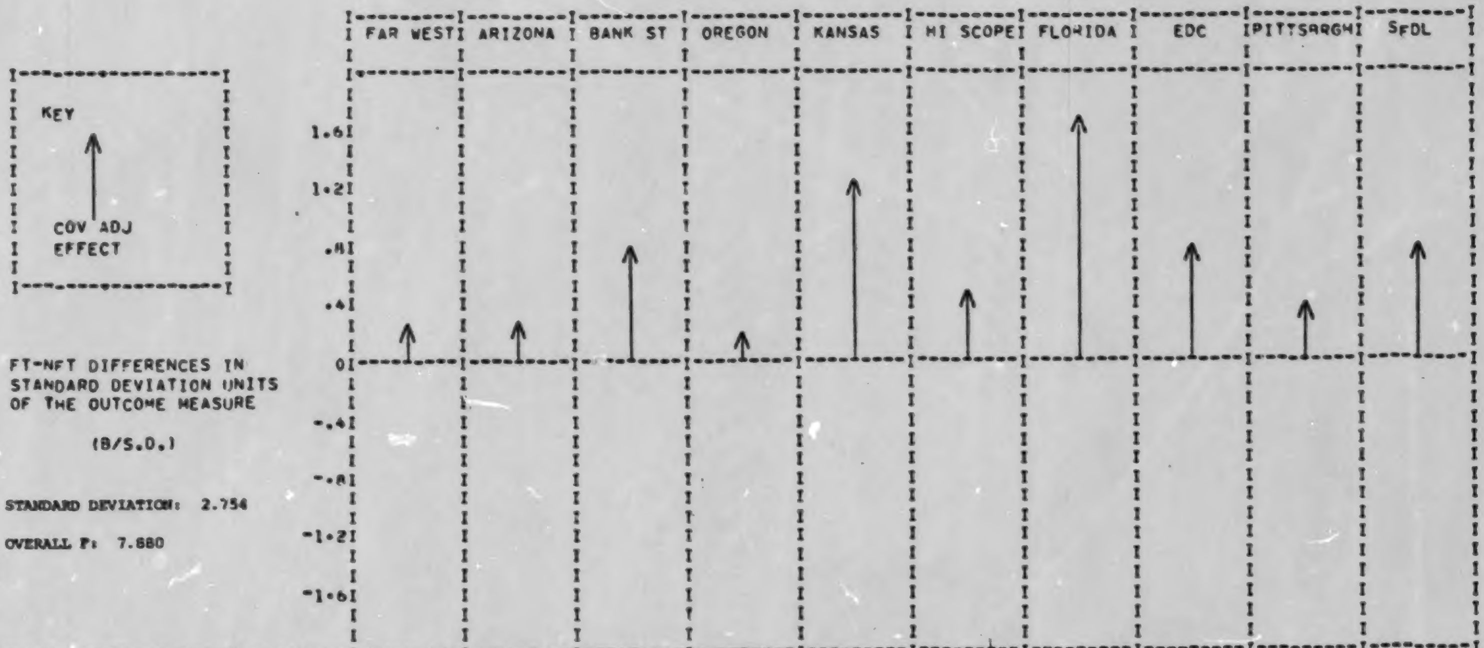
OVERALL F: 2.474

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SFDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		1.274	3.806	5.586	1.358	4.328	3.274	-1.281	-1.238	1.580	3.289
S.E. = STANDARD ERROR OF B		1.767	1.794	2.128	1.970	2.020	2.312	2.209	2.537	3.992	1.824
F= SIGNIFICANCE STATISTIC		.520	* 4.504	* 6.894	.475	* 4.591	2.006	.336	.238	.157	3.236
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		.171	.512	.751	.183	.582	.440	-.172	-.166	.212	.442
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

TABLE VIII-19

SPONSOR PROFILE FOR AMOUNT OF TEACHER TRAINING: PARENT INVOLVEMENT



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

(B/S.D.)

STANDARD DEVIATION: 2.754

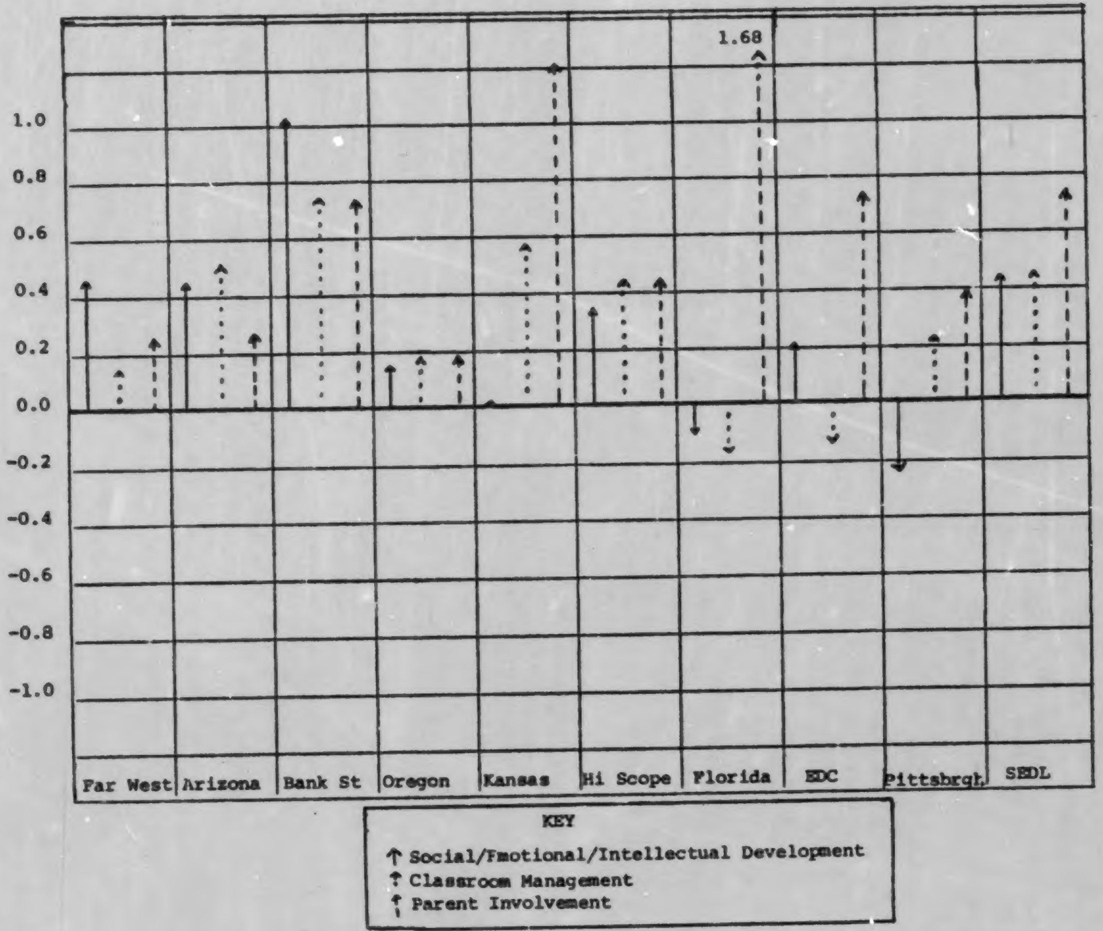
OVERALL F: 7.880

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SFDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		.760	.763	2.049	.545	3.310	1.208	4.630	2.060	1.092	2.022
S.E. = STANDARD ERROR OF B		.614	.623	.739	.684	.701	.803	.767	.881	1.386	.635
F= SIGNIFICANCE STATISTIC		1.534	1.502	*7.697	.634	*22.281	2.263	*36.423	*5.467	.620	*10.142
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		.276	.277	.744	.198	1.202	.439	1.681	.748	.397	.734
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

TABLE MII-20

SPONSOR PROFILES OF COVARIATE ADJUSTED SCORES IN STANDARD DEVIATION UNITS FOR THE AMOUNT OF TRAINING GIVEN IN THREE AREAS: SOCIAL/EMOTIONAL/INTELLECTUAL DEVELOPMENT, CLASSROOM MANAGEMENT, AND PARENT INVOLVEMENT



3.4 Reported Teacher Practices

Four cluster variables were developed for investigating the reported teacher practices in the following areas:

- Emphasis on Basic Skills vs. Affective Development,
- Flexible/Open Classroom,
- Individualization of Instruction, and
- Teacher Centeredness.

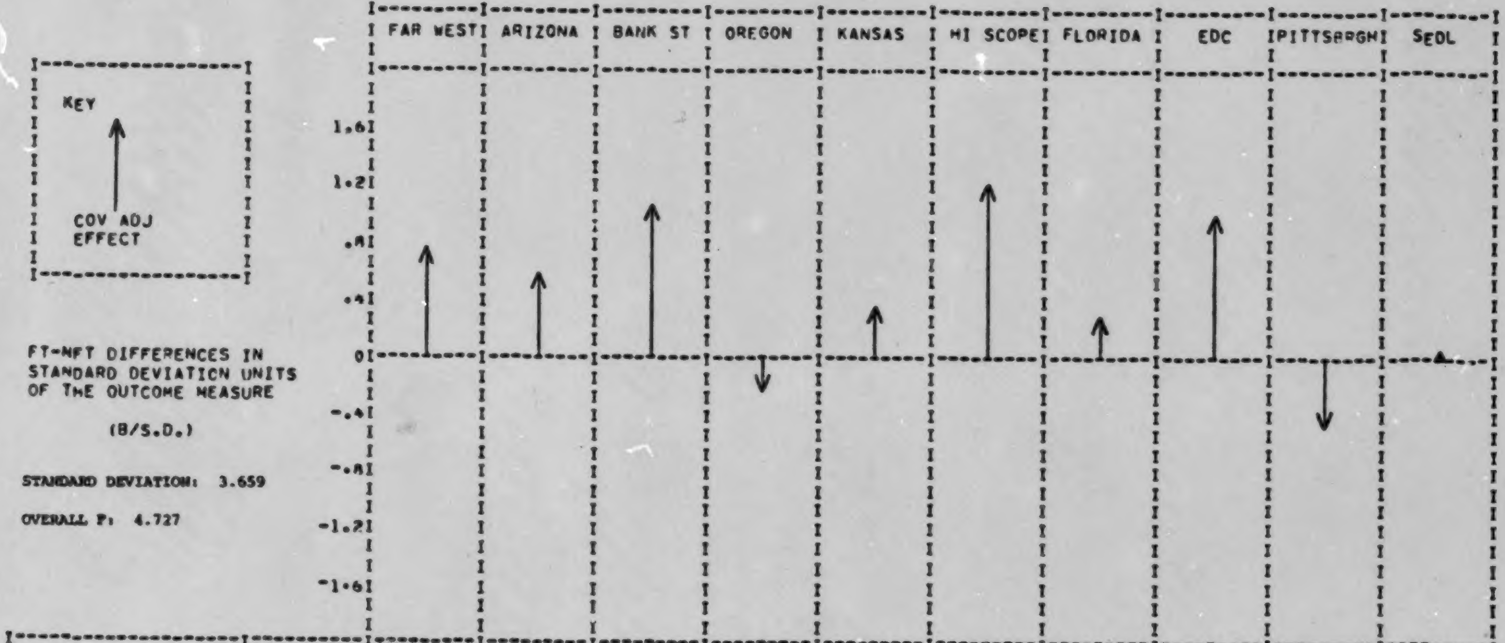
Tables 21 through 23 present the contrasts for the latter three clusters. The cluster, Emphasis on Basic Skills vs. Affective Development, did not contain an overall significant F and, therefore, was dropped from further consideration.

Table 21 shows the FT/NFT contrasts for reported teacher practices in using a flexible/open classroom. The FT teachers associated with Far West, Arizona, Bank Street, High/Scope, and EDC all report using this practice to a significantly greater extent than the NFT comparison groups. In general, all Sponsors with the exception of Oregon and Pittsburgh have teachers who employ this practice to a greater extent than the NFT groups. This is an anticipated pattern for both of these Sponsors.

Table 22 displays contrasts which indicate that Arizona, Bank Street, Kansas, High/Scope, and SEDL have FT teachers who practice individualization of instruction to a significantly greater extent than their NFT comparisons. Only Oregon and Pittsburgh are below their respective NFT groups on this dimension. The cluster is designed to focus on the simultaneous occurrence of many different kinds of activities. Pittsburgh's position on this table may reflect the practice of simultaneously scheduling different levels of the same curriculum activity rather than a lack of individualization of instruction, per se.

Table 23 indicates that there are considerable Sponsor differences on the dimension of teacher-centeredness. Far West, Bank Street, and High/Scope have FT teachers who are significantly below their NFT comparisons on this dimension. Oregon and Kansas, on the other hand, have FT teachers who are significantly above their NFT groups. FT teachers in Pittsburgh and SEDL also report some degree of teacher centeredness.

TABLE MII-21
SPONSOR PROFILE FOR REPORTED TEACHER PRACTICES: FLEXIBLE / OPEN CLASSROOM



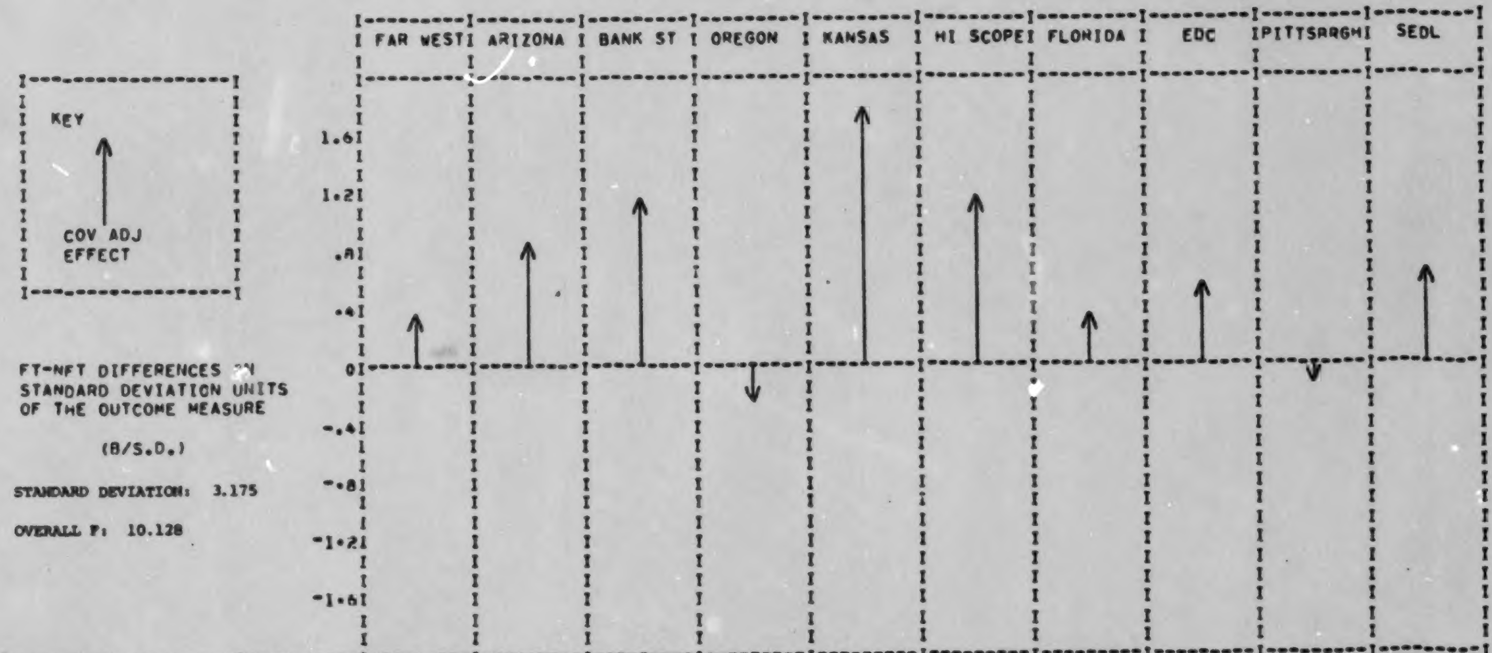
STANDARD DEVIATION: 3.659

OVERALL F: 4.727

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		2.716	2.004	3.725	-.799	1.135	4.345	.960	3.457	-1.803	.249
S.E. = STANDARD ERROR OF B		.817	.830	.984	.912	.934	1.010	1.022	1.174	1.847	.846
F= SIGNIFICANCE STATISTIC		* 11.079	* 5.832	* 14.320	.768	1.474	* 16.503	.882	* 8.676	.954	.087
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		.742	.548	1.018	-.218	.310	1.187	.262	.948	-.493	.068
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

TABLE VII-22
SPONSOR PROFILE FOR REPORTED TEACHER PRACTICES: INDIVIDUALIZATION OF INSTRUCTION



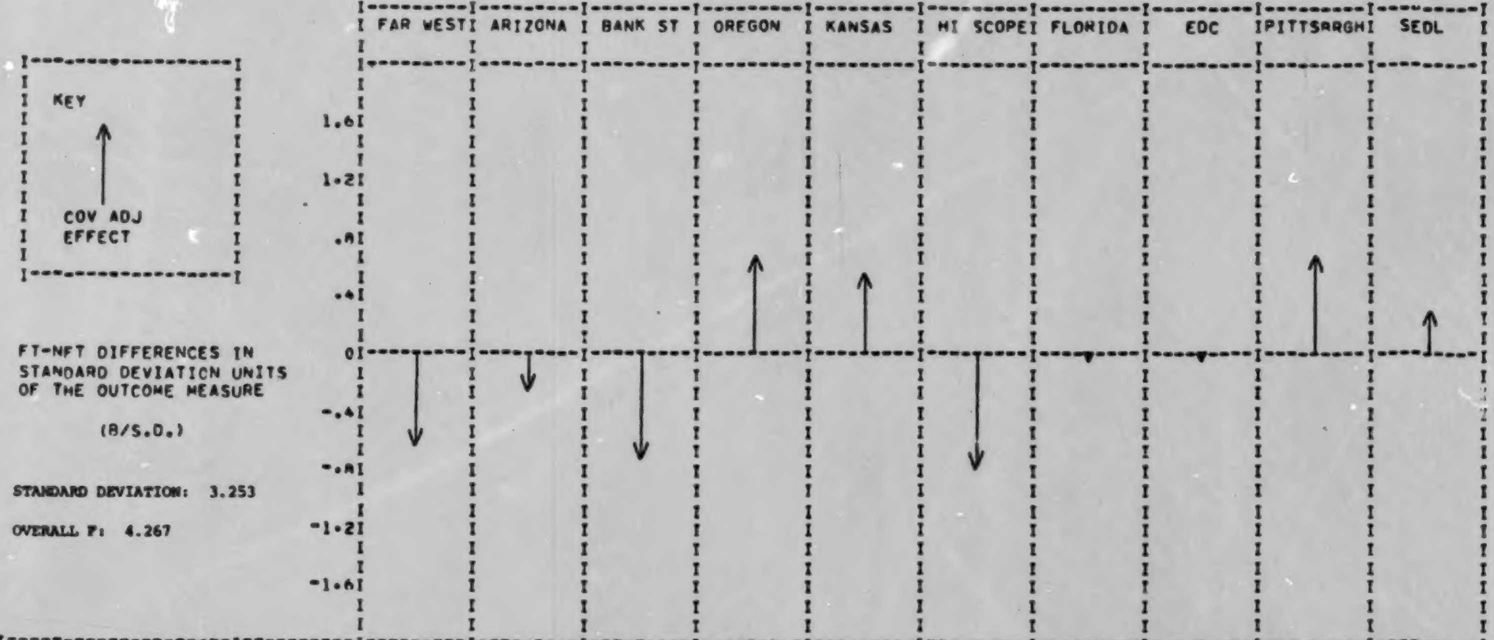
STANDARD DEVIATION: 3.175

OVERALL F: 10.128

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		1.206	2.648	3.584	-.783	5.556	3.724	.995	1.830	-.353	1.941
S.E. = STANDARD ERROR OF B		.685	.695	.825	.764	.783	.896	.857	.984	1.548	.709
F= SIGNIFICANCE STATISTIC		3.100	* 14.495	* 18.872	* 1.049	* 50.330	* 17.263	1.349	3.462	.060	* 7.496
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		.380	.834	1.129	-.247	1.750	1.182	.313	.576	-.111	.611
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

TABLE HII-23
SPONSOR PROFILE FOR REPORTED TEACHER PRACTICES: TEACHER CENTEREDNESS



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

(B/S.D.)

STANDARD DEVIATION: 3.253

OVERALL F: 4.267

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		-2.039	-.874	-2.319	2.242	1.763	-2.552	-.082	-.140	2.228	.912
S.E. = STANDARD ERROR OF B		.733	.744	.883	.818	.838	.959	.917	1.053	1.657	.759
F= SIGNIFICANCE STATISTIC		*7.728	1.378	*6.897	*7.519	*4.422	*7.079	.087	.018	1.810	1.445
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		-.627	-.269	-.712	.689	.542	-.785	-.025	-.043	.685	.280
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

Table 24 displays the profiles for the three reported teacher practices. The variability among the Sponsor groups is apparent, and in most cases, consistent with Sponsor models. For example, Far West, Arizona, Bank Street and High Scope show profiles which are consistent with their programs which emphasize flexible/open classrooms, differentiation in terms of simultaneous instructional activities, and lack of teacher-centeredness. Variations in the size as well as the direction of the contrasts are also relatively consistent with each Sponsor's approach. The Kansas profile indicates that the individualized instruction is at the teacher's direction. Further, the reward system with its component of free play during part of the day is depicted as a comparatively open classroom by these FT teachers. Oregon is shown as a structured classroom, involved in largely group work, where the teacher initiates most of the learning and playing activities. Pittsburgh shows a similar pattern, with more emphasis on adult supervision and choice of learning materials. The absence of an upward arrow for individualization of instruction was discussed previously. EDC is shown as an open classroom with individualized instruction in comparison to its NFT group. Florida shows similarities to Sponsor 11. The SEDL teachers indicate a comparatively high degree of individualized instruction which is teacher initiated and supervised.

3.5 Parent Involvement

The responses of the teachers to aspects of the importance of parent involvement were constructed into four cluster variables:

- Social/educational activities,
- Decision making activities,
- Housekeeping activities, and
- Benefit of parent involvement.

All FT teachers in all Sponsor groups placed greater importance on parent involvement in social and educational activities than did their respective NFTs. Therefore, the profile for this cluster variable has been omitted from further discussion.

Table 25 shows the contrasts between the FT and the NFT teachers' opinions about involving parents in decision making activities. High/Scope, Florida, and Pittsburgh have FT teachers who value this aspect of parent

TABLE III-24
SPONSOR PROFILES OF COVARIATE ADJUSTED SCORES
IN STANDARD DEVIATION UNITS FOR REPORTED TEACHER PRACTICES:
FLEXIBLE OPEN CLASSROOM, INDIVIDUALIZATION OF INSTRUCTION, TEACHER CENTEREDNESS

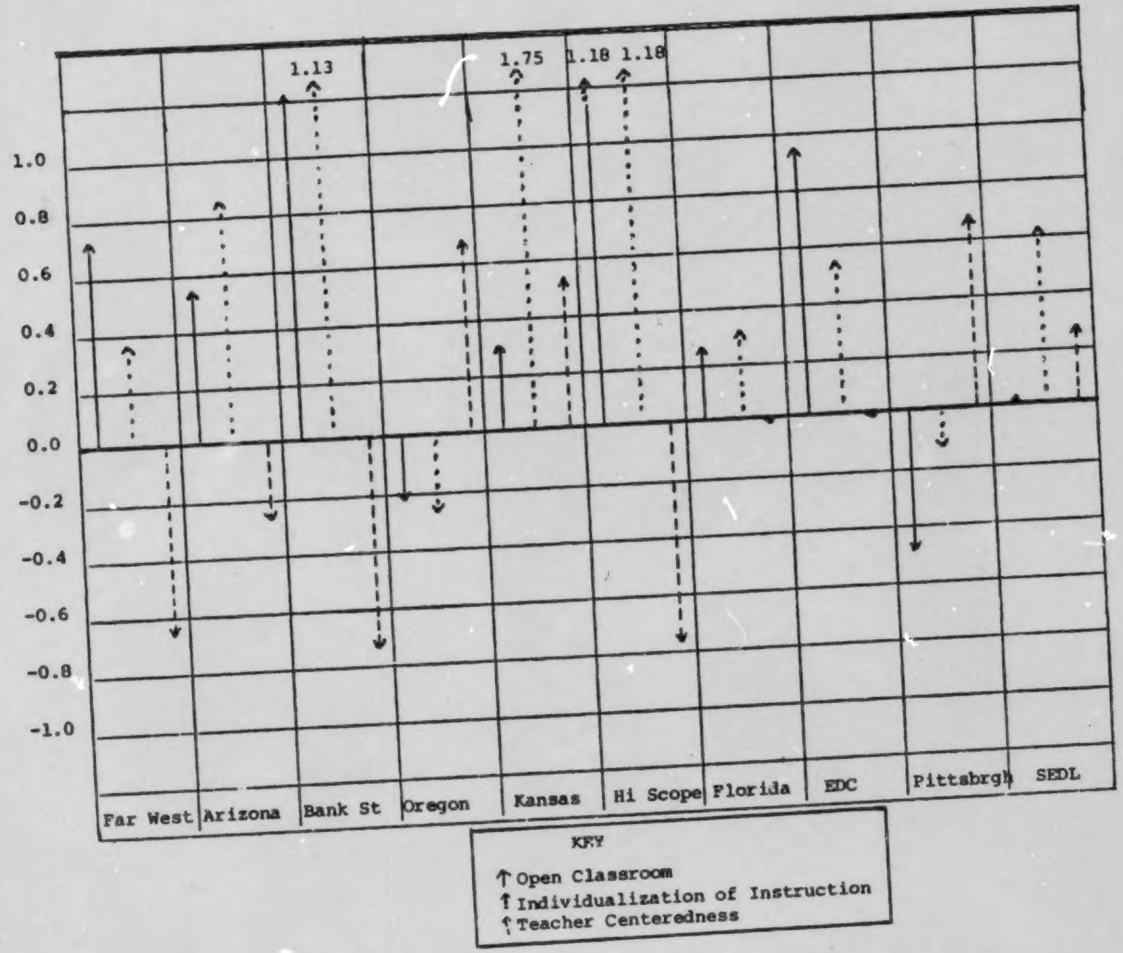
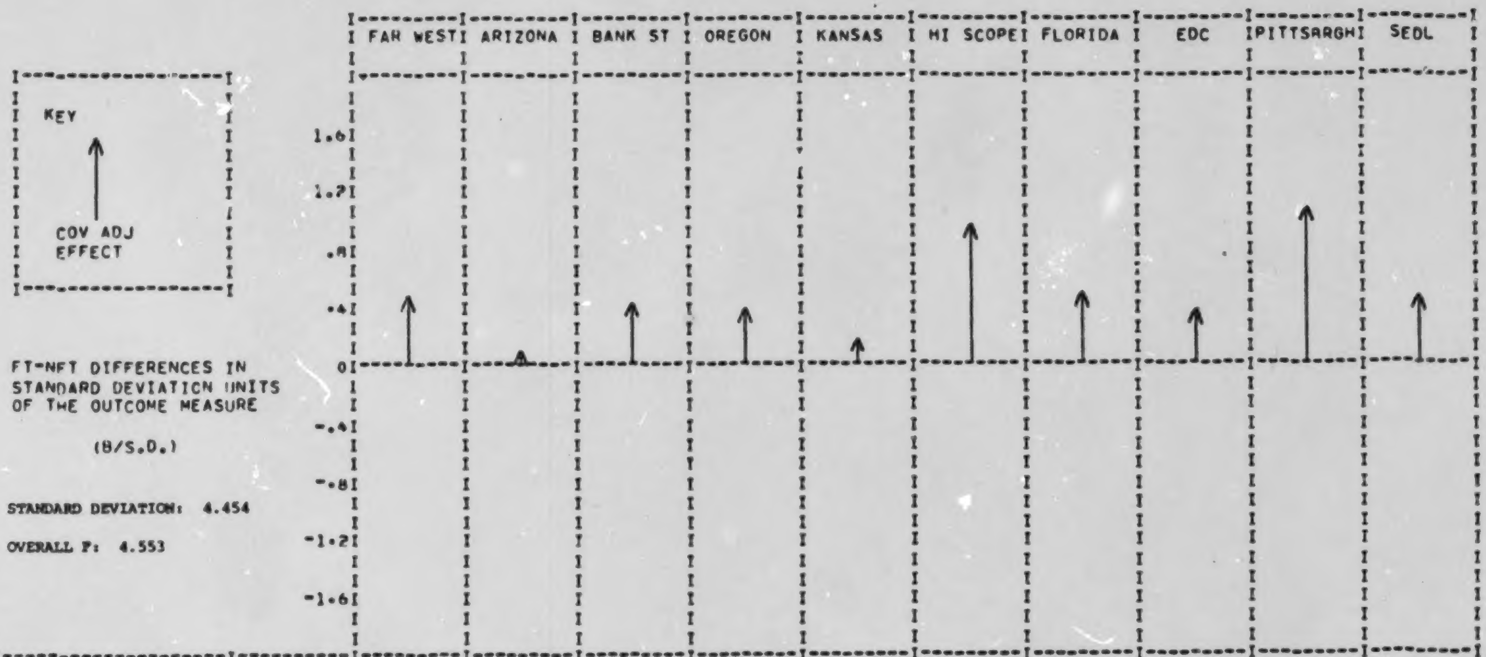


TABLE MII-25

SPONSOR PROFILE FOR IMPORTANCE OF PARENT INVOLVEMENT: PARENT DECISION MAKING ACTIVITIES



STANDARD DEVIATION: 4.454

OVERALL F: 4.553

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		1.499	.428	1.780	1.685	.685	4.317	2.132	1.723	4.519	2.016
S.E. = STANDARD ERROR OF B		1.025	1.040	1.234	1.143	1.171	1.341	1.281	1.471	2.315	1.060
F= SIGNIFICANCE STATISTIC		2.139	.169	2.080	2.174	.342	*10.371	*2.769	1.372	*3.810	3.615
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		.480	.096	.400	.378	.154	.969	.479	.387	1.015	.453
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

involvement to a far greater extent than their NFT groups. In general, all FT teachers value parent involvement in decision making to a higher degree than their comparison groups.

Table 26 contains only one significant contrast. This is based on reports by FT teachers associated with Oregon who indicate they do not use parents for housekeeping activities. Only Kansas, EDC and SEDL appear to value parent involvement in such activities to any degree.

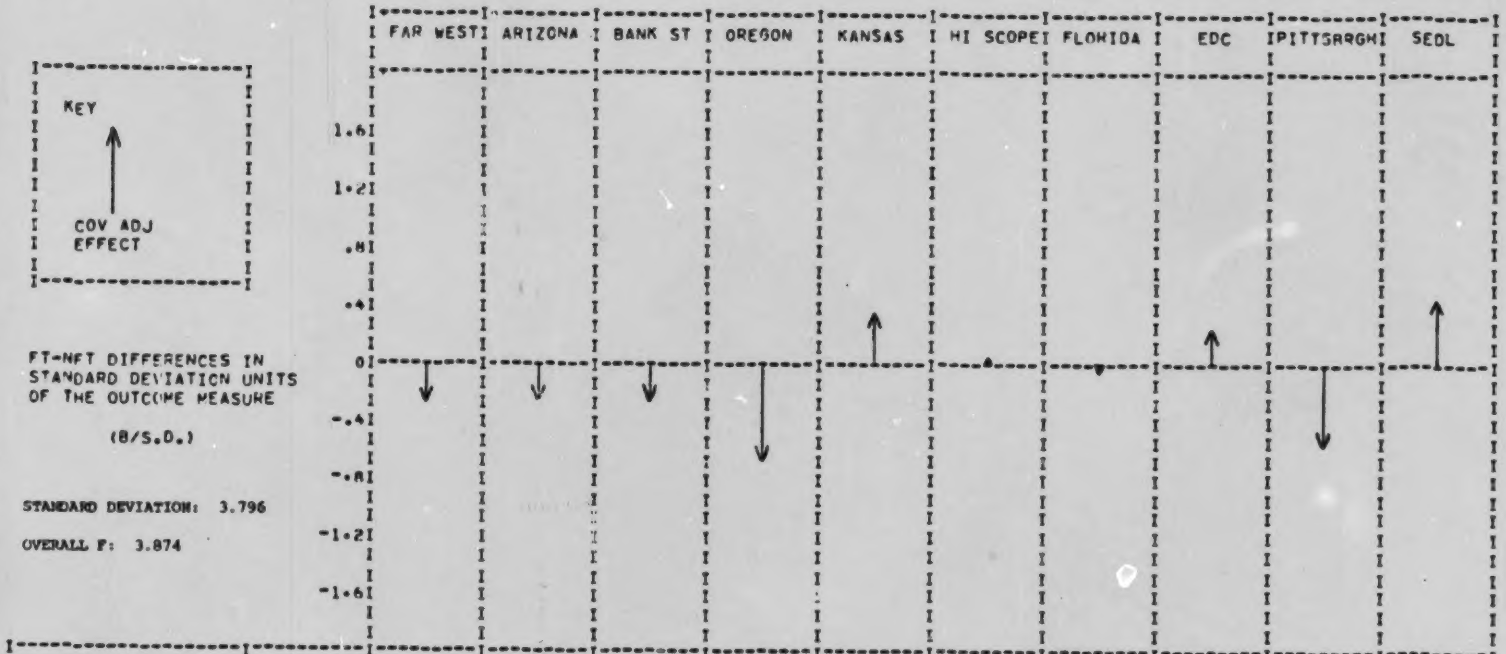
The contrasts in Table 27 indicate the teachers' opinions on the benefits that they, the children, or the program have received from parent involvement. There are five significant Sponsor contrasts in which the FT teachers indicate more benefit from parent involvement than the NFT comparisons. These are Far West, Bank Street, Kansas, Florida and SEDL. Only Arizona, Oregon, and High/Scope have NFT teachers who identify a greater benefit from parent involvement than the FT groups.

The profiles of the parent involvement clusters are presented in Table 28. It is apparent that all Sponsors' FT teachers express opinions which indicate a greater support for various aspects of parent involvement than shown by the NFT groups. Both High/Scope and Pittsburgh have NFT comparison groups which involve parents in decision making activities to a very small extent compared with other NFT groups. This tends to elevate those contrasts above those of Florida, a model which is based upon parent involvement. The Florida profile suggests that the FT teachers using this approach make a relatively sharp distinction between involving parents in decision making and involving parents with the pupil's learning process at home. In general, FT appears to be involving parents in decision making to some degree. Further, most FT groups identify more benefit from parent involvement than the NFT groups. This may be indicative of an important new dimension being systematically incorporated into the educational process.

Table 29 indicates the relative amounts of PAC activity among the various Sponsors. It appears that Bank Street has the most active PAC's and Pittsburgh has the least active PAC groups. Florida, a parent educational model, is located near the mean of all Sponsors. It does not appear that there is a direct correlation between the amount of PAC activity and the importance teachers place on involving parents in the decision making process.

TABLE HII-26

SPONSOR PROFILE FOR IMPORTANCE OF PARENT INVOLVEMENT: HOUSEKEEPING ACTIVITIES



STANDARD DEVIATION: 3.796

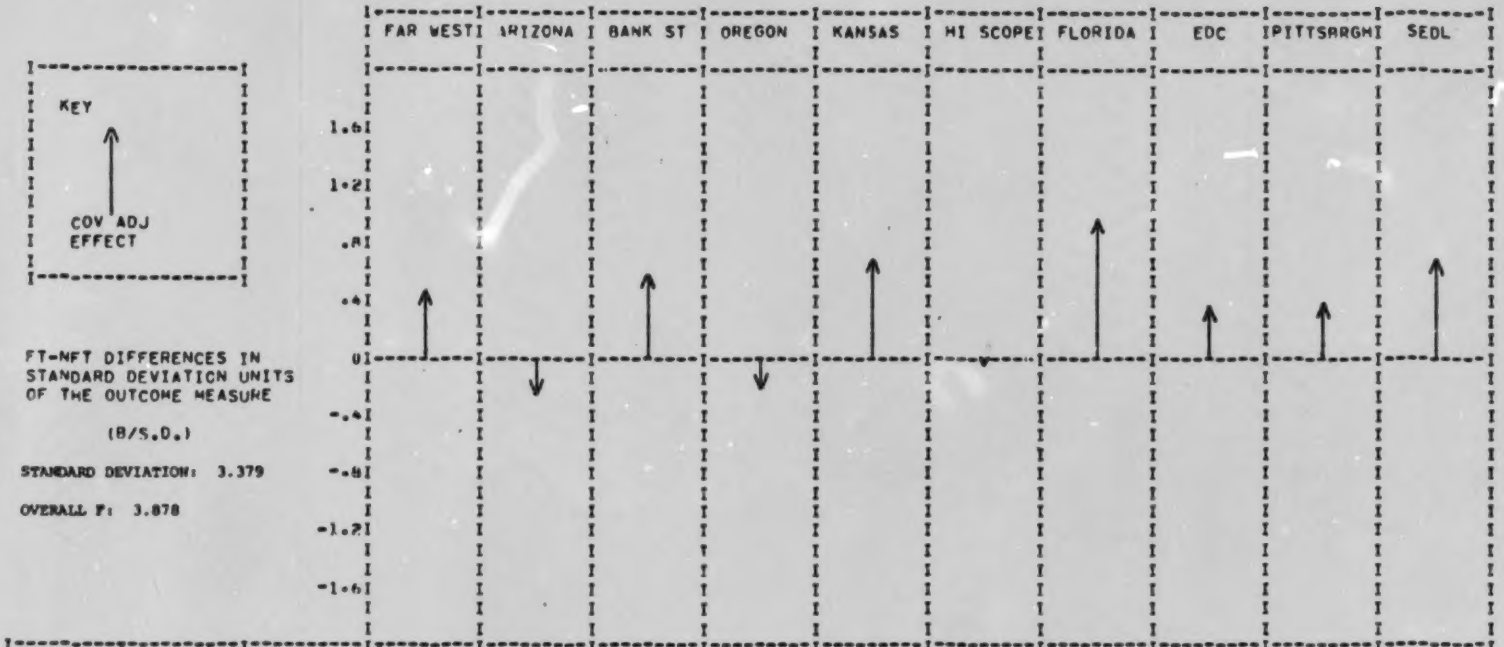
OVERALL F: 3.874

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		-.955	-.960	-1.000	-2.492	1.253	.072	-.016	.846	-2.060	1.733
S.E. = STANDARD ERROR OF B		.874	.887	1.053	.975	.999	1.440	1.093	1.255	1.975	.905
F= SIGNIFICANCE STATISTIC		1.192	1.170	.931	* 6.531	1.572	.004	0.000	.455	1.088	3.669
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		-.252	-.253	-.263	-.656	.330	.019	-.004	.223	-.543	.457
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

TABLE MII-27

SPONSOR PROFILE FOR BENEFIT OF PARENT INVOLVEMENT



STANDARD DEVIATION: 3.379

OVERALL F: 3.878

		FAR WEST	ARIZONA	BANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		1.593	-.976	1.993	-.704	2.094	-.210	3.098	1.126	1.339	2.360
S.E. = STANDARD ERROR OF B		.786	.798	.947	.877	.899	1.029	.983	1.129	1.777	.814
F= SIGNIFICANCE STATISTIC		* 4.104	1.495	* 4.431	.644	* 5.426	0.000	* 9.929	.994	.568	* 8.408
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		.471	-.289	.590	-.208	.620	-.062	.917	.333	.396	.698
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

TABLE MII-28
SPONSOR PROFILES OF COVARIATE ADJUSTED SCORES IN STANDARD DEVIATION UNITS FOR IMPORTANCE OF PARENT INVOLVEMENT: DECISION-MAKING ACTIVITIES, HOUSEKEEPING ACTIVITIES, AND BENEFIT OF PARENT INVOLVEMENT

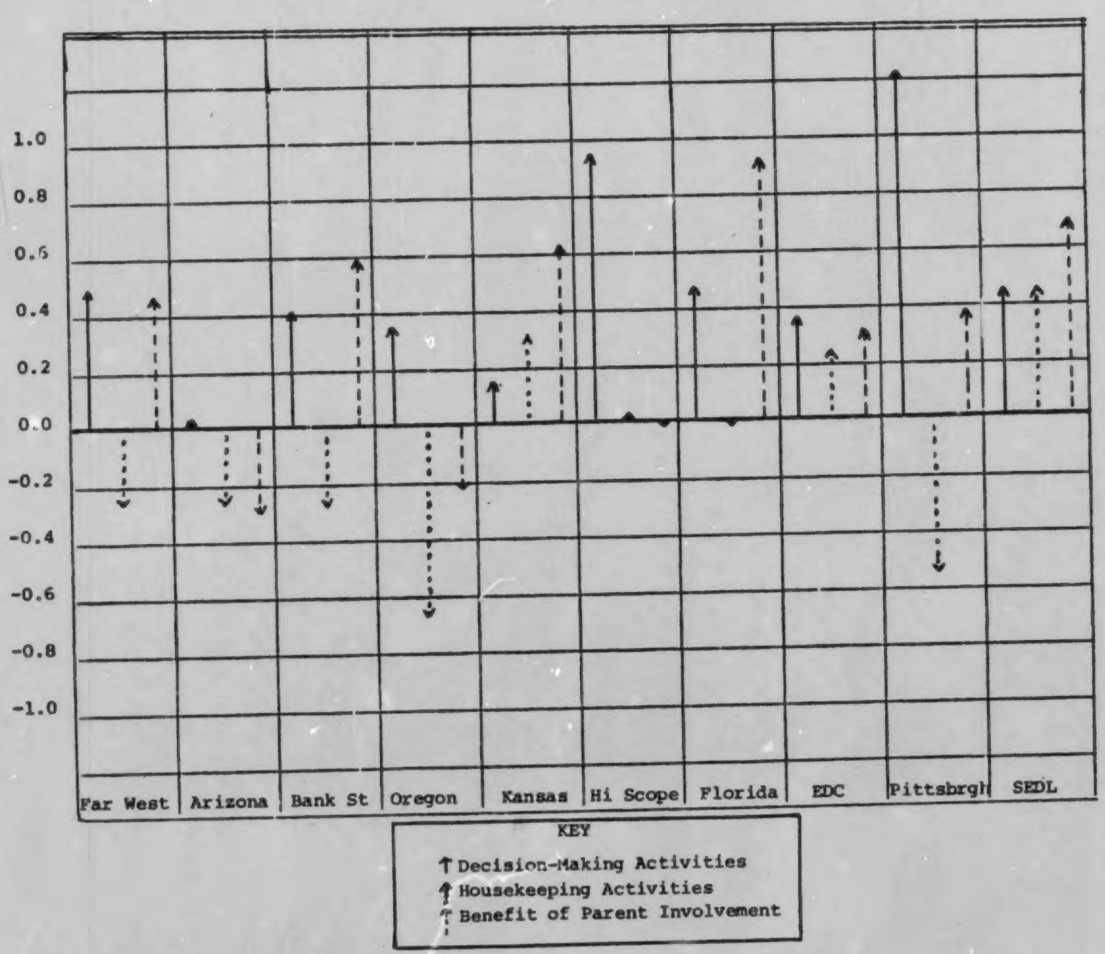


TABLE MII-29
MEANS AND STANDARD DEVIATIONS FOR AMOUNT OF PAC ACTIVITY FOR FT TEACHERS BY SPONSOR AND COHORT/GRADE

SPONSOR	Kindergarten, Cohort IV FT N	 $\overline{X}$	 SD	Non-entering First, Cohort III FT N	 $\overline{X}$	 SD	Other* FT N	 $\overline{X}$	 SD
Far West	22	20.955	8.352	42	17.690	7.377	18	16.222	7.109
Arizona	23	15.217	6.742	44	18.023	6.983	8	18.125	6.854
Bank Street	19	23.526	6.785	23	23.609	6.535	43	18.442	6.794
Oregon	3	17.333	6.429	34	19.618	6.135	31	19.161	7.326
Kansas	26	21.923	6.157	43	20.465	6.497	30	18.533	7.080
High Scope	10	20.800	7.300	27	21.630	8.039	16	18.813	6.483
Florida	6	22.333	5.428	29	19.379	6.894	28	19.036	7.753
EDC	12	18.667	5.914	19	15.947	6.786	32	19.281	5.390
Pittsburgh	13	17.077	8.558	5	12.000	4.848	11	16.000	4.450
SEDL	7	12.000	5.099	35	18.114	7.588	-	-	-
TOTALS	141	19.518	7.492	301	19.166	7.191	217	18.465	6.718

*Includes teachers in Cohort III, second grade; Cohort II, second and third grades; and Cohort I, third grade.

3.6 Teacher Satisfaction with Pupils' Progress

One cluster variable, teacher satisfaction with pupils' affective development and two items concerning the teachers' satisfaction with progress in math and in reading were used to investigate this teacher attitude. Only one variable, teacher satisfaction with progress in reading, produced a significant overall F test.

As seen in Table 30, only three Sponsors' FT teachers report relatively more satisfaction with their pupils: Arizona, High/Scope, and Pittsburgh. These three Sponsors have FT teachers who are also generally satisfied with the model. The FT teachers associated with Far West and Florida, on the other hand, are not as satisfied with their pupils' reading progress as the NFT teachers are although Florida teachers report satisfaction with the model in general.

4.0 CONCLUSIONS

The Follow Through program does appear to be having a measurable impact on its teachers. They are receiving more training than their comparisons in social/emotional/intellectual development, classroom management, and parent involvement. The teacher practices which are being employed vary across Sponsors in terms of the use of an open/flexible classroom, individualization of instruction, and teacher-centeredness. These variations are, generally, consistent with each Sponsor's approach. FT also appears to be having a considerable impact on the use of parent involvement and the teachers' attitudes toward its use. These teachers' reports indicate that the parents' potential contributions are being sought and utilized to an increasing extent over that found in NFT groups. Finally, most FT teachers are satisfied with their Sponsor's model. This implies that the Sponsors have designed programs which are generally amenable to teachers' styles and pupils' needs for learning. While the match may not be perfect, most teachers appear to have adopted their Sponsor's approach to the education process.

While we would have liked to have merged teacher and pupil data for this report, time and contractual limitations prevented us from doing so. While the teacher studies give us some confidence that we are in fact assessing the impact of different models, we do not yet know the extent to which variations in implementation affect pupil outcomes.

TABLE III-30

SPONSOR PROFILE FOR TEACHER SATISFACTION WITH PUPILS' PROGRESS IN READING

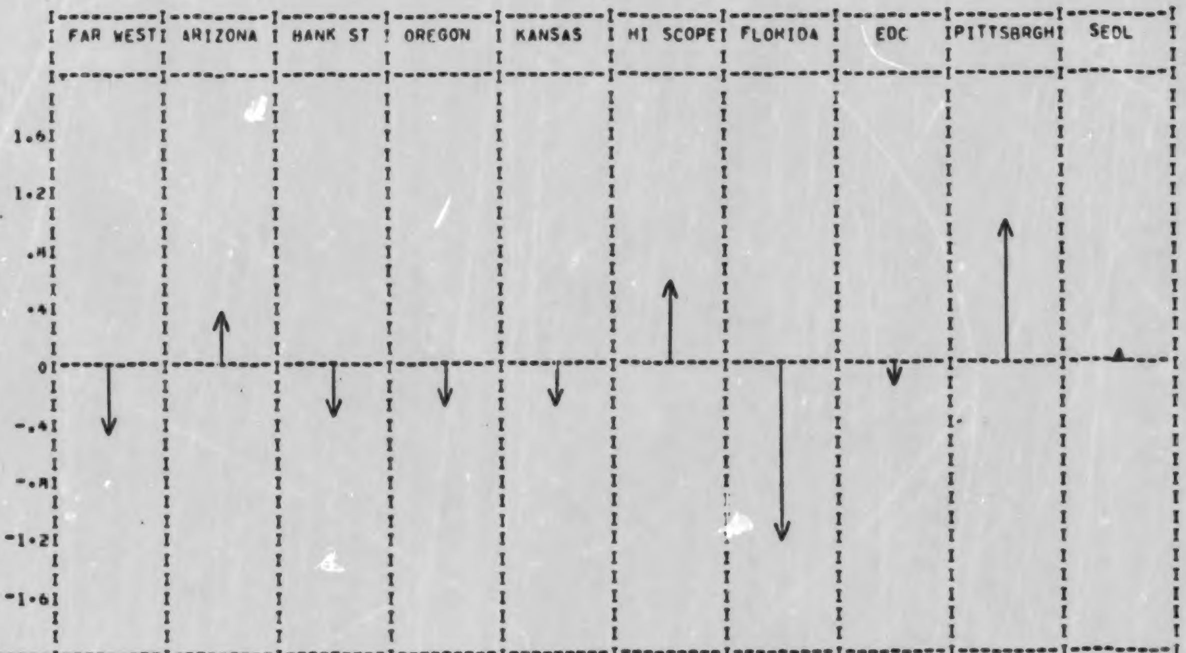
KEY
COV ADJ
EFFECT

FT-NFT DIFFERENCES IN
STANDARD DEVIATION UNITS
OF THE OUTCOME MEASURE

(B/S.D.)

STANDARD DEVIATION: 0.864

OVERALL F: 3.423



		FAR WEST	ARIZONA	HANK ST	OREGON	KANSAS	HI SCOPE	FLORIDA	EDC	PITTSBRGH	SEDL
B=SPONSOR EFFECT (FT-NFT DIFFERENCES)		-.422	.287	-.296	-.207	-.258	.474	-.889	-.150	.832	.002
S.E. = STANDARD ERROR OF B		.197	.200	.237	.219	.225	.257	.246	.282	.444	.203
F= SIGNIFICANCE STATISTIC		* 4.603	2.067	1.564	.895	1.320	3.395	* 13.075	.283	3.517	.006
B/SD = EFFECT IN SD UNITS (SEE GRAPH)		-.488	.332	-.343	-.240	-.299	.549	-1.029	-.173	.963	.002
N	FT	39	35	19	25	30	21	26	13	10	30
	NFT	31	33	31	31	24	19	19	23	05	34

*p less than .05

MONOGRAPH III

SCHOOL ENTRY

MONOGRAPH III

SCHOOL STUDY

1.0 INTRODUCTION

The several layers of factors which affect pupil performance have been briefly set forth in Volume IIA.
There, pupil characteristics, family background charactertistics, class variables (both teacher related and peer group related), school variables and community characteristics are all discussed. In this Monograph, we focus on the latter two spheres of influences--school and community charactertistics--and describe the FT/NFT groups (overall and within Sponsor) in terms of several contextual variables. In order to truly assess the effects of the various FT programs, it is important to know the environments in which both they and their comparison groups function.

The data for this study were drawn from three sources:

1. Information collected by the Technical Consultants assigned by The U.S. Office of Education to each participating local educational agency. These consultants provided technical support to each agency through systematic visits during which information on the properties of each school was collected. The checklist (School Questionnaire) used by these consultants appears on Table 1.
2. Directory of Elementary and Secondary Schools in Selected Districts: Enforllment and Staff by Racial Ethnicity Group. U.S. Office of Civil Rights, 1971.
3. The Fourth Count: School District Census. U.S. Bureau of the Census, 1970.[1]

2.0 METHOD

2.1 The Sample

Technical Consultants collected information from those schools in which FT/NFT pupils were tested during the 1971-1972 school year.

---

[1] The Fourth Count School District Census aggregates U.S. census data (1970) to school district level.

TABLE MIII-1

SCHOOL QUESTIONNAIRE

School Information

(fill out one for each Follow Through and Comparison School)

District: ______________________ **City, State:** ______________________

Follow Through Project: ______________________ **School:** ______________________

All questions refer to the 1971-72 school year (last year) unless otherwise specified.

1. Which grade levels were taught in this school last year? (CHECK ALL THAT APPLY. If school is ungraded, indicate what the equivalent grade levels of the pupils in the school would be according to the childrens' ages.)

Grade level	Pre-K	K	1	2	3	4	5	6	7	8	9	10	11	12
Check those in this school														

2. Estimate the total enrollment of the school as of October 1, 1971 ______________

In the following chart, fill in the number of Follow Through and non-Follow Through classes and the number of pupils in those classes by grade level. (If ungraded, indicate what the equivalent grade levels of the classes and pupils would be according to the childrens' ages.)

3.

Grade	Follow Through		Non-Follow Through	
	Classes	Pupils	Classes	Pupils
K				
1				
2				
3				

4. How many of the children noted above in the grades K through 3 received food free or at a reduced price (that is, at a price less than the current, locally set fooo price)?

	Number of Follow Through Children (K–3)	Number of Non-Follow Through Children (K–3)
Breakfast	______________	______________
Lunch	______________	______________
Milk	______________	______________
Snack	______________	______________

5. Which of the following best describes the location of this school?

_____ 1. Large city, over 500,000 people
_____ 2. Large city, over 200,000 to 500,000
_____ 3. Suburb of a large city
_____ 4. Rural area near a large city
_____ 5. Middle-sized city, 50,000 to 200,000
_____ 6. Suburb of a middle-sized city
_____ 7. Rural area near a middle-sized city
_____ 8. Small city or town, less than 50,000 population
_____ 9. Rural area, not near a large or middle-sized city

6. Listed below are a number of special services available to pupils at various schools. If any of these are available in your school, please estimate the average number of man-days per week allocated for each service. Estimate to the nearest half of a man-day. Give separate estimates for all pupils in the school and for FT pupils. (e.g. If 2 mathematics specialists serve all pupils in the school 2 days a week each, mark 4 (man-days) in column 1. If one of these man-days is reserved for Follow Through pupils only, mark 1 (man-day) in column 2.)

Persons Providing the Service	Average number of man-days allocated for all pupils in this school, including Follow Through	Average number man-days allocated for Follow Through pupils only
1. Guidance counselor		
2. Child welfare and attendance worker		
3. Psychologist/psychiatrist		
4. Physician (other than psychiatrist)		
5. Nurse		
6. Speech teacher/therapist		
7. Social worker		
8. Reading specialist		
9. Math specialist		
10. Music/art teacher		
11. Physical Education Teacher		
12. Librarian		
13. Audiovisual specialist		
14. Curriculum specialist/resource teacher/teacher trainer		
15. English as a second language/bilingual consultant		
16. Community liaison person		
17. Other (specify)		

Directions for Completing Item 6

Listed in Item 6 are a number of special services available to pupils at various schools. If any of these are available in your school, please estimate the average number of man-days per week allocated for each service. Estimate to the nearest half of a man-day as follows:

1. In column 1, enter the number of man-days per week allocated to Follow Through pupils, but not to non-Follow Through pupils.
2. In column 2, enter the number of man-days per week allocated to non-Follow Through pupils, but not to Follow Through pupils.
3. In column 3, enter the number of man-days per week allocated to any child in the school on the basis of need, regardless of the child's program assignment.
4. In column 4, enter the total number of man-days of service per week available to pupils in this school. This number should equal the sum of column 1, 2, and 3.

EXAMPLES

School X has the following persons providing services to its pupils:

1. Two half-time guidance counselors assigned to the school, 2 1/2 days each. One counselor serves Follow Through pupils only. The other counselor services non-Follow Through pupils only.
2. One full-time nurse who serves any child in the school who needs medical attention, five days per week.
3. One reading specialist who serves non-Follow Through pupils only, 1 day per week.
4. One ESL consultant, assigned 3 1/2 days per week. Two days are allocated for Follow Through pupils only and 1 1/2 days are allocated for any child who needs assistance.
5. Three full-time community liason persons serving Follow Through pupils only, five days a week each.

	Number of Man Days Per Week Allocated for Each Service			
Persons Providing the Service	1 Follow Through pupils only	2 Non-Follow Through pupils only	3 Any pupil	4 Total 1+2+3
1. Guidance counselor	2 1/2	2 1/2	0	5
2. Nurse	0	0	5	5
3. Reading specialist	0	1	0	1
4. ESL consultant	2	0	1 1/2	3 1/2
5. Community liaison person	15	0	0	15

Item 6

Please read the instructions and examples on the reverse side carefully before responding.

	Number of Man Days Per Week Allocated for Each Service			
Persons Providing the Service	1 Follow Through pupils only	2 Non-Follow Through pupils only	3 Any pupil	4 Total 1+2+3
1. Guidance counselor				
2. Child welfare and attendance worker				
3. Psychologist/psychiatrist				
4. Physician (other than psychiatrist)				
5. Nurse				
6. Speech teacher/therapist				
7. Social worker				
8. Reading specialist				
9. Math specialist				
10. Music/art teacher				
11. Physical Education Teacher				
12. Librarian				
13. Audiovisual specialist				
14. Curriculum specialist/resource teacher/teacher trainer				
15. English as a second language/ bilingual consultant				
16. Community liaison person				
17. Other (specify)				

7. For the 1971-72 school year, excluding summer school,

a. Did this school receive Title 1 funds?

_____ 1. Yes _____ 2. No _____ 3. Don't Know

If yes, estimate the amount of money this school received under Title 1 to the nearest thousand. $ __________

b. Did this school receive Follow Through funds?

_____ 1. Yes _____ 2. No _____ 3. Don't Know

If yes, estimate the amount of money this school received under Follow Through to the nearest thousand. $ __________

c. Did this school receive funds from other federal programs (e.g. ESEA, TITLE II, III, VI, VII, or VIII; NDEA III or VA; Civil Rights Act, TITLE IV, Vocational Education Act; Education Professions Development Act)?

_____ 1. Yes _____ 2. No _____ 3. Don't Know

If yes, estimate the amount of money this school received under other federal programs to the nearest thousand. $ __________

d. Estimate the total budget for this school including federal monies to the nearest thousand. $ __________

8. Listed below are a number of reasons why various schools have been assigned to either the Follow Through group or the comparison group. Check the reasons this school was assigned to either Follow Through or comparison status.

_____ 1. All schools in this district are Follow Through (comparison) schools

_____ 2. All schools in this district were randomly assigned to Follow Through (comparison) status

_____ 3. This school had a high concentration of children from low income, minority, and/or bilingual families

_____ 4. This school had a mixture of children from families of different economic, ethnic, or racial backgrounds

_____ 5. This school had a high concentration of children with Head Start or other preschool experience

_____ 6. This school was eligible for Title I and was receiving Title I funds

_____ 7. This school was eligible for Title I but was not receiving Title I funds

_____ 8. Other (specify) ____________________

THE LAST TWO QUESTIONS APPLY ONLY TO SCHOOLS WHICH HAVE FOLLOW THROUGH CLASSES

9. If a child leaves a FT class, is a new pupil assigned to fill his place?

_________ 1. Yes _________ 2. No

If "yes", how is the replacement selected? (check all that apply)

_________ 1. By randomly selecting a replacement from among all children in the same grade level not in FT

_________ 2. There is no established basis for selection

_________ 3. By selecting from among those children with previous Head Start or pre-school experience

_________ 4. By selecting from among those who have had FT experience in another school or district

_________ 5. By selecting from among those who come from low-income, minority, or non-English speaking families

10. How were teachers assigned to FT classes in this school? (check all that apply)

_________ 1. Teachers were randomly assigned from all available teachers at the appropriate grade levels

_________ 2. There was no established basis for assignment

_________ 3. Teachers asked to teach in FT

_________ 4. Only teachers with specialized education and training were selected

_________ 5. Only teachers with previous teaching experience were selected

_________ 6. Parents screened and helped select FT teachers

_________ 7. FT teachers were assigned after other positions were filled

_________ 8. Other (specify) ____________________________

This school study includes schools associated with one of the ten Sponsors included in the pupil analyses. In total, there are 229 FT schools and 207 NFT schools analyzed in the present study. A breakdown by Sponsor is displayed in Table 2.

The number of schools available for description varies with the variable being studied because complete information was not available for each school. Further, some schools were not listed in the Office of Civil Rights Directory because of the sampling criteria for small- and medium-sized districts. Finally, one school district was not included in the Fourth Count Census Tapes.

2.2 Variables

From these sources, several variables were developed to describe:

- school size and structure,
- school funds,
- school services, and
- school district characteristics.

Total school enrollment, pupil/teacher ratio, entry grade, and highest grade in school provide a general description of school size and structure. These variables were developed from the School Questionnaire, except the pupil/teacher ratio which was taken from the Office of Civil Rights Directory.

Data on school funds were also collected by means of the School Questionnaire. Although we do not know exactly how many children are eligible for Title I services, the ratio of school funds to total number of children in the school does provide useful information. If a school has few Title I dollars per pupil, we can assume that either (a) there are very few Title I children in the school on which these dollars are concentrated or (b) the school has a reasonable number of Title I eligible children, but receives only a negligible amount of funds. In a school with a high ratio of Title I funds to the number of children in the school, we can safely assume that there are a substantial number of Title I children who are benefiting from a sizable amount of Title I dollars.

The presence and amount of other State or Federal funds are also examined. Although we do not know the exact source for these funds, they do suggest that certain schools are receiving a sizeable amount of funds in addition to Title I and/or Follow Through.

TABLE MIII-2

NUMBER OF FT AND NFT SCHOOLS FOR EACH SPONSOR

Sponsor	Number of FT Schools	Number of NFT Schools
Far West	29	27
Arizona	34	37
Bank Street	32	25
Oregon	25	28
Kansas	23	18
High/Scope	14	13
Florida	32	20
EDC	17	17
Pittsburgh	12	14
SEDL	11	8
TOTAL	229	207

School service information was the major concern of the School Questionnaire. For descriptive purposes, the services have been separated into professional and non-professional categories. The professional services include the following:

- Medical--doctor, nurse;
- Social--social worker, child welfare worker;
- Instructional--reading specialist, math specialist, curriculum specialist, speech therapist, English as a second language specialist;
- Activities--music and/or art teacher, physical education teacher; and
- Library-Av--librarian, audiovisual specialist.
- Psychological

The non-professional category includes community liaison persons and other paraprofessionals such as nurses' aides, reading aides, and social work interns.

For each personnel category specified in the Questionnaire, the allocation of services in mandays per week was indicated for three categories of pupils: FT pupils, NFT pupils, and any pupil in the school. Five mandays per week represents the equivalent of one full-time person. A ratio of mandays of service per week to the number of children eligible for that service was computed to estimate the availability of services for each target group. An FT service score, therefore, is composed of the ratio of the mandays of service per week available to FT children to the number of FT children in the school plus the ratio of that service for any pupil to the total number of children in the school, as in the formula below:

$$\begin{array}{c}\text{Mandays per week}\\\text{FT service}\end{array} = \frac{\begin{array}{c}\text{Mandays per week}\\\text{for FT pupils only}\end{array}}{\text{Number of FT pupils}} + \frac{\begin{array}{c}\text{Mandays per week}\\\text{for any pupil}\end{array}}{\text{Total enrollment}}$$

The score for NFT is similarly composed:

$$\begin{array}{c}\text{Mandays per week}\\\text{NFT service}\end{array} = \frac{\begin{array}{c}\text{Mandays per week}\\\text{for NFT pupils only}\end{array}}{\text{Number of NFT pupils}} + \frac{\begin{array}{c}\text{Mandays per week}\\\text{for any pupil}\end{array}}{\text{Total enrollment}}$$

In the tables and figures which follow, we have converted mandays of service per week per child to a slightly different metric for purposes of convenience and ease of understanding. Services are reported in number of full-time or full time equivalent professionals available per hundred children where full time is equal to five mandays per week. This metric is intended to give the reader a better perspective on the relative caseloads of the various service personnel.

The school district characteristics discussed in this chapter are as follows:

- labor force characteristics, including
  - percentage of persons 16 years or over in each of the occupational categories defined by the census
  - percentage of females in the labor force
- income level, including
  - mean income
  - percentage of families below poverty level
- language, including
  - percentage of non-English speaking persons
  - percentage of Spanish speaking persons
- education level, defined as mean years of education for persons 25 years and over
- family structure, defined as percentage of families with female head of household and children under 18
- mobility, defined as percentage of persons who have always lived in the same house

These district characteristics will give us a general picture of the environment in which the schools operate.

2.3 Analytic Techniques

Given the wide variability of the school size, pupil/teacher ratio, funding and service variables, frequency distributions were employed to display each variable by FT/NFT and Sponsor. In most cases, the frequency distributions for the total population were divided into five equal parts with approximately 20 percent of the population in each category. In the cases of medical, psychological, and non-professional services, where zero is a well represented category, it is also displayed. Means and standard deviations are reported for the school district variables.

3.0 RESULTS

3.1 School Size and Structure

Figures 1 through 4 illustrate that there is very little difference between FT and NFT schools in entry grade, highest grade in school, or total enrollment. FT has slightly more schools with prekindergarten as the entry grade than NFT as seen in Figure 1. This is consistent with the FT objective of building on preschool experience. Having prekindergarten programs in the same schools with the FT programs may bring about closer articulation between the two. FT also has slightly more schools where the highest grade is three (see Figure 2), again reflecting the emphasis of the program on the primary grades. The majority of schools in this study, though, have kindergarten as the entry grade and six as the highest.

The total enrollment (Figure 3) is fairly evenly distributed across categories for both FT and NFT, with most of the schools having from 251 to 550 children. FT has slightly more schools, however, in the 551-850 range.

The pupil/teacher ratio, as illustrated in Figure 4, indicates an FT/NFT difference; the FT schools tend to have lower pupil/teacher ratios than the NFT schools. This, no doubt, reflects the FT guidelines for limiting the number of pupils in each class.

Tables 3-6 display these variables by Sponsor. As can be seen in Table 3 Far West has the largest percentage of schools with prekindergarten as the entering grade. This is consistent for both FT (52%) and NFT (44%). Furthermore, there are no schools with first grade as the entering grade. Pittsburgh has the second largest percentage of FT schools with prekindergarten; however, few Pittsburgh NFT schools have preschool classes. Kansas also has a large percentage (30.4%) of FT schools with prekindergarten classes, but none of the NFT schools have them. None of the Kansas schools have first grade as the entering grade.

As can be seen in Table 4, SEDL has the largest percentage of FT and NFT schools where a primary grade (K, 1, 2, 3) is the highest grade in the school (45.4% and 50% respectively). Oregon has a large percentage (41.7%) of FT schools where a primary grade is the highest grade, but only 7.7 percent of the NFT schools are also structured in this way. Moreover,

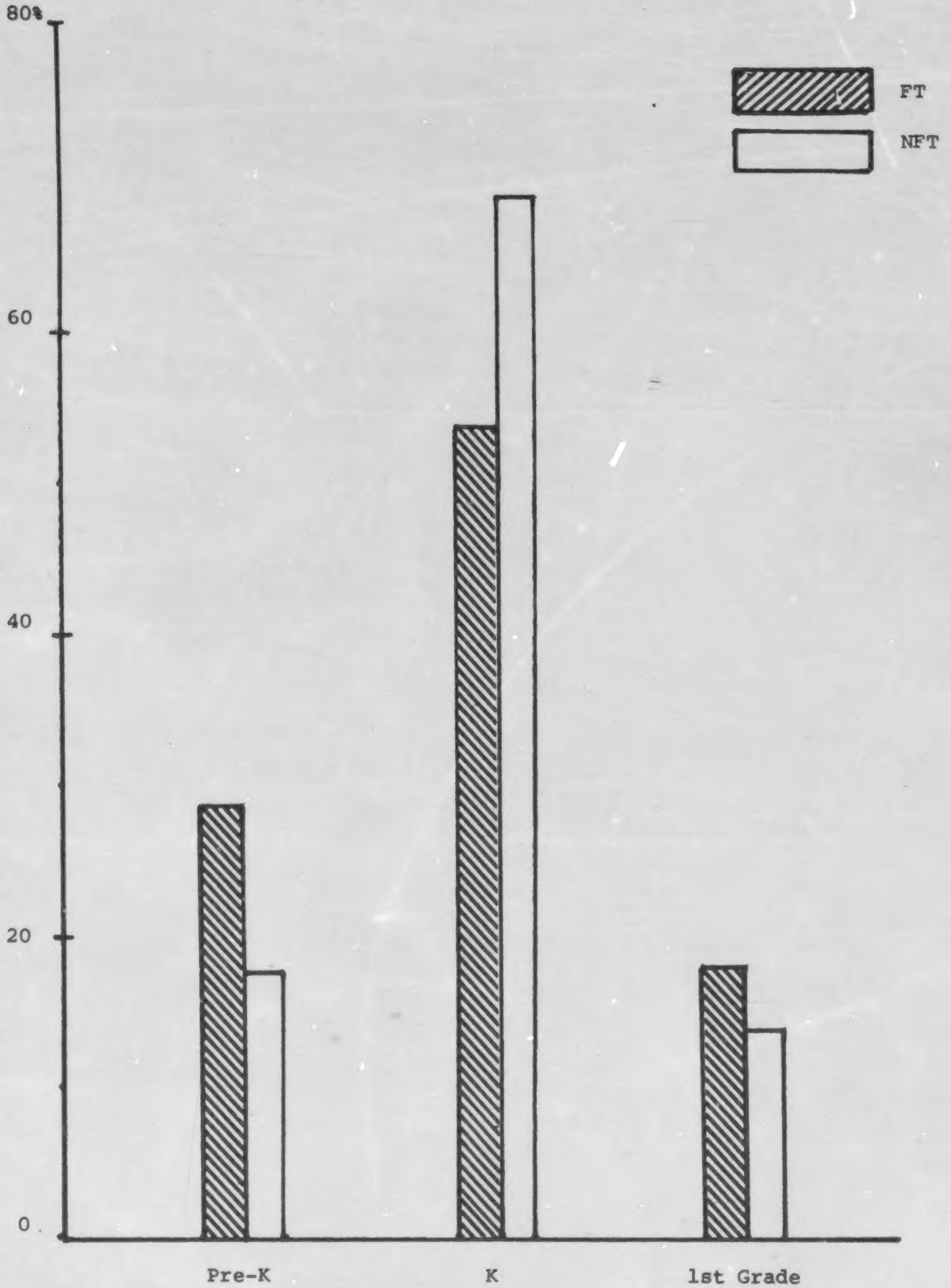


Figure MIII-1. Frequency Distribution of Entry Grade For Follow Through and Non-Follow Through

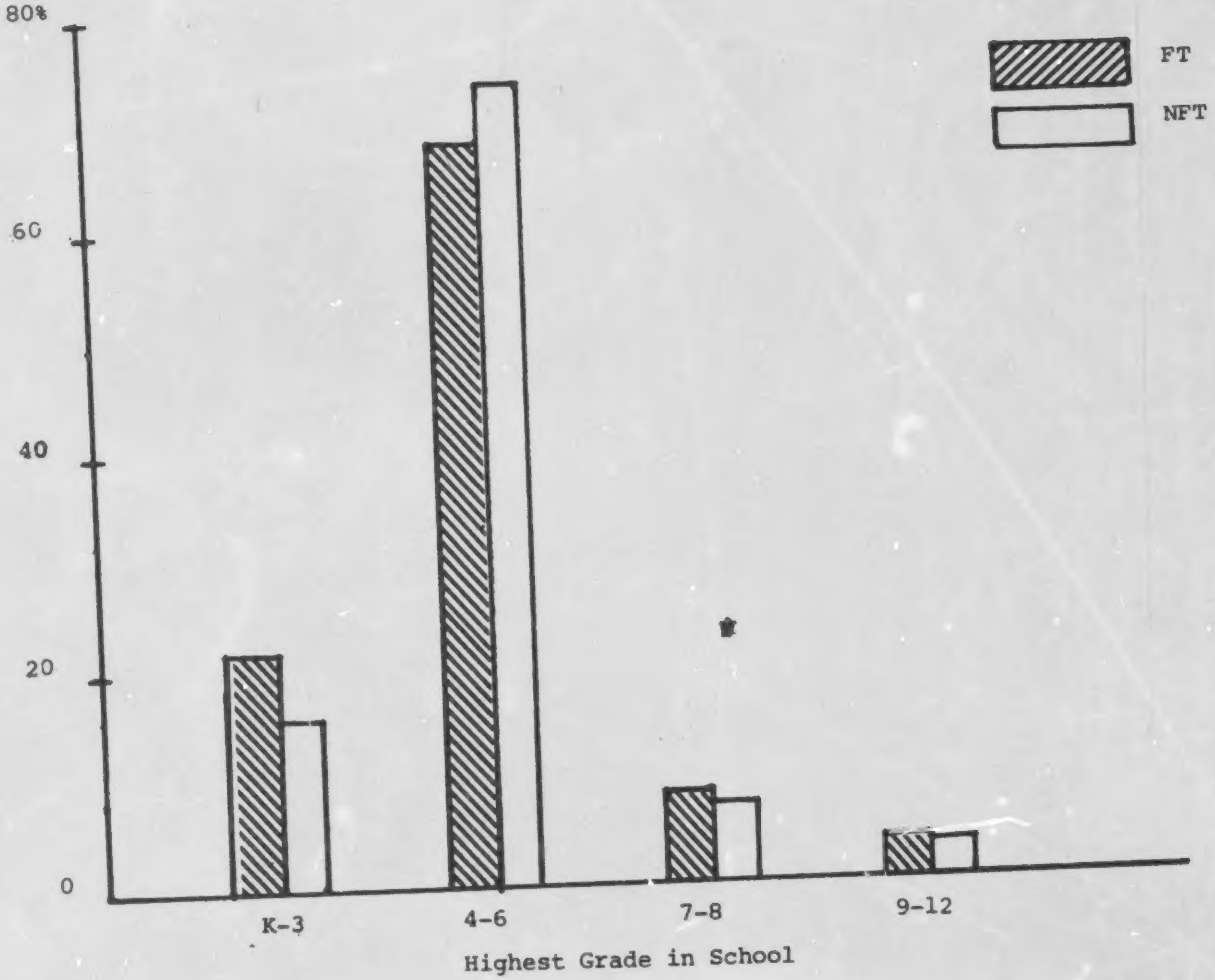


Figure MIII-2. Frequency Distribution of Highest Grade in School for Follow Through and Non-Follow Through

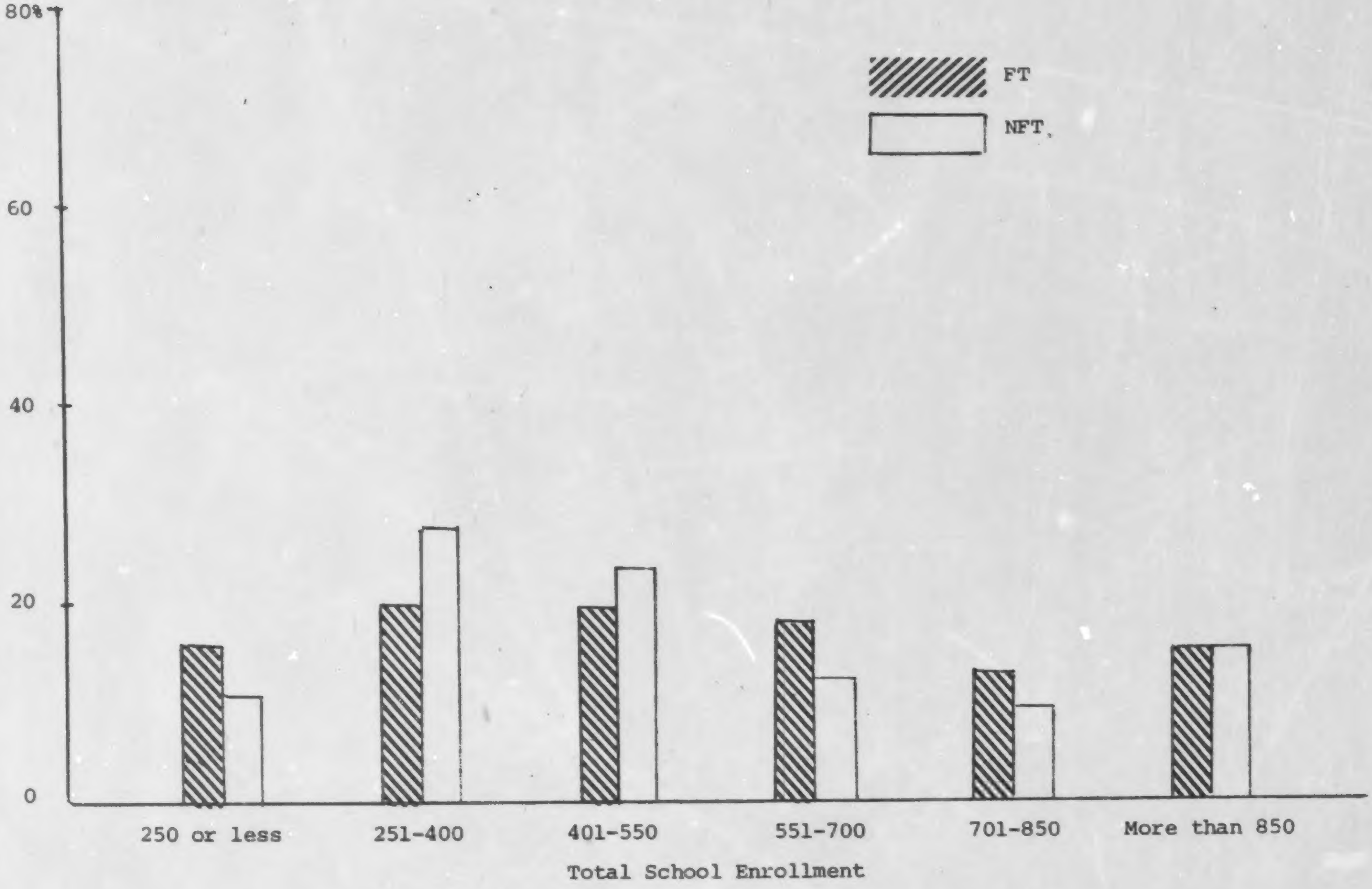


Figure MIII-3. Frequency Distributions of Total School Enrollment for Follow Through and Non-Follow Through

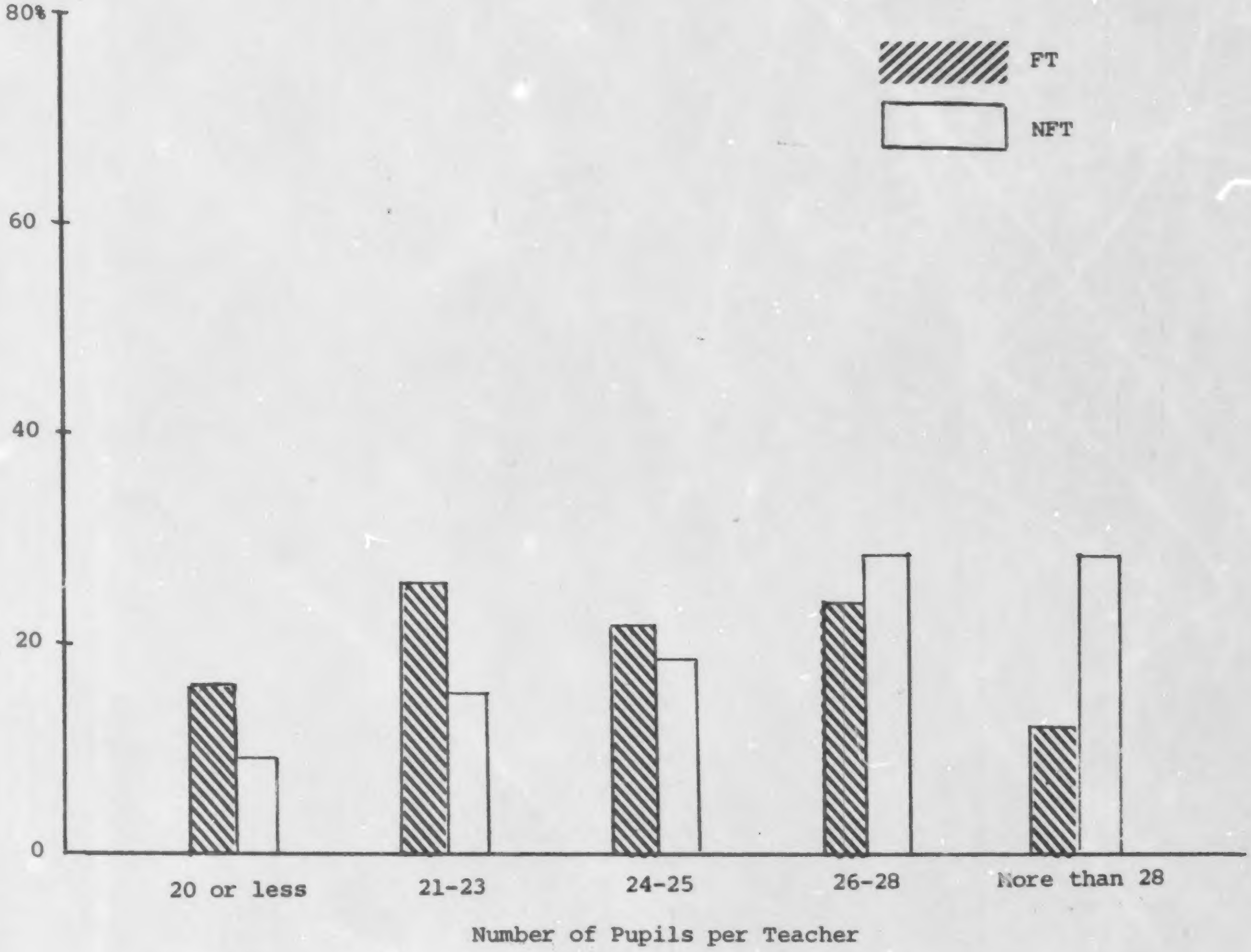


Figure MIII-4. Frequency Distributions of Pupil-Teacher Ratio for Follow Through and Non-Follow Through

TABLE MIII-3

FREQUENCY COUNTS OF
ENTERING GRADE IN SCHOOL
FOR EACH SPONSOR'S FT AND NFT GROUP

		Entry Grade					
		FT			NFT		
Sponsor		PreK	K	1st	PreK	K	1st
Far West	N	15	14	0	12	15	0
	%	51.7	48.3	0	44.4	55.6	0
Arizona	N	7	17	10	5	24	7
	%	20.6	50.0	29.4	13.9	66.7	19.4
Bank Street	N	8	16	8	6	17	2
	%	25.0	50.0	25.0	24.0	68.0	8.0
Oregon	N	4	18	2	3	17	6
	%	16.7	75.0	8.3	11.5	65.4	23.1
Kansas	N	7	16	0	0	18	0
	%	30.4	69.6	0	0	100.0	0
High Scope	N	4	4	6	3	5	5
	%	28.6	28.6	42.9	23.1	38.5	38.5
Florida	N	6	13	10	2	15	3
	%	20.7	44.8	34.5	10.0	75.0	15.0
EDC	N	4	10	3	1	12	4
	%	23.5	58.8	17.6	5.9	70.6	23.5
Pittsburgh	N	5	6	1	2	10	2
	%	41.7	50.0	8.3	14.3	71.4	14.3
SEDL	N	4	6	1	2	6	0
	%	36.4	54.5	9.1	25.0	75.0	0

TABLE MIII-4

FREQUENCY COUNTS OF HIGHEST GRADE IN SCHOOL FOR EACH SPONSOR'S FT AND NFT GROUP

		Highest Grade in School							
		FT				NFT			
Sponsor		K,1,2,3	4,5,6	7,8	9,10, 11,12	K,1,2,3	4,5,6	7,8	9,10 11,12
Far West	N	5	19	5	0	7	14	6	0
	%	17.2	65.5	17.2	0	25.9	51.8	22.2	0
Arizona	N	9	24	1	0	7	28	0	1
	%	26.5	70.6	2.9	0	19.4	77.8	0	2.8
Bank Street	N	3	22	4	3	0	22	3	0
	%	9.4	68.7	12.5	9.4	0	88.8	12.0	0
Oregon	N	10	14	0	0	2	21	2	1
	%	41.7	58.3	0	0	7.7	80.8	7.7	3.8
Kansas	N	3	16	2	1	2	13	1	1
	%	13.6	72.7	9.1	4.5	11.8	76.4	5.9	5.9
High Scope	N	1	8	1	4	0	12	0	1
	%	7.1	57.1	7.1	28.6	0	92.3	0	7.7
Florida	N	5	24	0	0	4	17	0	0
	%	17.2	82.8	0	0	20.0	80.0	0	0
EDC	N	2	10	5	0	2	11	4	0
	%	11.8	58.8	29.4	0	11.8	64.7	23.5	0
Pittsburgh	N	4	8	0	0	1	11	0	2
	%	33.3	66.7	0	0	7.1	78.6	0	14.3
SEDL	N	5	4	2	0	4	4	0	0
	%	45.4	36.4	18.2	0	50.0	50.0	0	0

while none of the FT schools for Oregon span the eighth or twelfth grade, 11.5 percent of the NFT schools do. High/Scope has a relatively large percentage of FT schools (28.6%) that span the high school grades, and few (7.1%) primary schools. None of the Florida schools span the junior high or high school grades. For the other Sponsors, the majority of the schools have kindergarten as the entering grade and 4,5,or 6 as the highest grade.

The total enrollment for the Sponsor FT and NFT groups varies widely, as seen in Table 5. In the cases of Bank Street, High/Scope, Florida and EDC, the FT schools are generally larger than their respective NFT schools. EDC has the most schools with large enrollments; 41.2 percent of FT schools have more than 850 pupils. Far West, Oregon, and Pittsburgh FT schools are smaller than their NFT schools. Approximately 42 percent of FT schools have fewer than 250 pupils. Arizona, Kansas, and SEDL FT/NFT schools are about equal in size.

Table 6 illustrates the FT/NFT by Sponsor differences in pupil/teacher ratio. Most of the Sponsors' FT schools have smaller ratios than their NFT schools. High/Scope has the largest percentage (42.9) of FT schools with small pupil/teacher ratios (23 or fewer pupils per teacher). Florida, on the other hand, has a substantial portion of its FT schools with large pupil/teacher ratios (40% have a ratio of 28 or more pupils per teacher), but the NFT schools are similarly composed. SEDL also has about equal ratios for FT and NFT. EDC is the only Sponsor where the FT ratio is larger than the NFT.

3.2 School Funds

Most of the schools in this study receive ESEA, Title I, and other Federal or State funds. As shown in Table 6, 92.5 percent of FT schools receive Title I and 88.6 percent receive other Federal or State funds. The percentage of NFT schools receiving Title I is 85.1, and the percentage receiving other Federal or State funds is 84.6. Overall, FT schools receive more Title I and other Federal or State dollars per pupil than NFT. Figure 5 shows that a higher percentage (22.5%) of FT schools than NFT schools (11.6%) receive in excess of $150 per pupil of Title I funds. This most likely reflects not more funds per pupil, but more children eligible for Title I in the schools receiving these dollars. As illustrated in Figure 7, the majority of both FT and NFT

TABLE MIII-5

FREQUENCY COUNTS OF TOTAL SCHOOL ENROLLMENT FOR EACH SPONSOR

		FT						NFT					
		Total School Enrollment						Total School Enrollment					
Sponsor		less than 250	251-400	401-550	551-700	701-850	more than 850	less than 250	251-400	401-550	551-700	701-850	more than 850
Far West	N	2	10	6	7	2	2	1	7	8	3	2	6
	%	6.9	34.5	20.7	24.1	6.9	6.9	3.7	25.9	29.6	11.1	7.4	22.0
Arizona	N	7	7	6	4	6	2	6	12	6	4	1	6
	%	21.9	21.9	18.8	12.5	18.8	6.3	17.1	34.3	17.1	11.4	2.9	17.1
Bank Street	N	7	4	8	6	3	4	3	9	4	5	3	1
	%	21.9	12.5	25.0	18.8	9.4	12.5	12.0	36.0	16.0	20.0	12.0	4.0
Oregon	N	9	7	3	4	2	0	3	12	6	5	1	1
	%	36.0	28.0	12.0	16.0	8.0	0	10.7	42.9	21.4	17.9	3.6	3.6
Kansas	N	2	1	6	4	3	7	2	1	5	3	2	5
	%	8.7	4.3	26[illegible]	17.4	13.0	30.4	11.1	5.6	27.8	16.7	11.1	27.8
High Scope	N	1	1	3	3	2	4	1	3	3	1	3	2
	%	7.1	7.1	21.4	21.4	14.3	28.6	7.7	23.1	23.1	7.7	23.1	15.4
Florida	N	0	8	9	8	4	3	1	5	9	2	1	1
	%	0	25.0	28.1	25.0	12.5	9.4	5.3	26.3	47.4	10.5	5.3	5.3
EDC	N	4	1	1	2	2	7	2	3	5	1	1	5
	%	23.5	5.9	5.9	11.8	11.8	41.2	11.8	17.6	29.4	5.9	5.9	29.4
Pittsburgh	N	5	4	1	0	1	1	3	6	1	2	1	1
	%	41.7	33.3	8.3	0	8.3	8.3	21.4	42.9	7.1	14.3	7.1	7.1
SEDL	N	1	1	1	1	3	4	1	1	1	0	2	3
	%	9.1	9.1	9.1	9.1	27.3	36.4	12.5	12.5	12.5	0	25.0	37.5

TABLE MIII-6

FREQUENCY COUNTS OF PUPIL/TEACHER RATIO FOR EACH SPONSOR

		PT					NFT				
		Number of Pupils Per Teacher					Number of Pupils Per Teacher				
Sponsor		20 or less	21-23	24-25	26-28	29 or more	20 or less	21-23	24-25	26-28	29 or more
Far West	N	8	6	4	7	3	4	2	3	7	10
	%	28.6	21.4	14.3	25.0	10.7	15.4	7.7	11.5	26.9	38.5
Arizona	N	4	12	7	9	1	3	5	5	11	11
	%	12.1	36.4	21.2	27.3	3.0	8.6	14.3	14.3	31.4	31.4
Bank Street	N	4	9	8	7	3	1	7	5	7	5
	%	12.9	29.0	25.8	22.6	9.7	4.0	28.0	20.0	28.0	20.0
Oregon	N	7	7	5	3	0	3	5	5	8	3
	%	31.8	31.8	22.7	13.6	0	12.5	20.8	20.8	33.3	12.5
Kansas	N	4	7	4	5	0	3	3	5	3	4
	%	20.0	35.0	20.0	25.0	0	16.7	16.7	27.8	16.7	22.2
High Scope	N	6	6	1	0	1	2	1	2	6	2
	%	42.9	42.9	7.1	0	7.1	15.4	7.7	15.4	46.2	15.4
Florida	N	0	2	7	9	12	0	4	3	4	8
	%	0	6.7	23.3	30.0	40.0	0	21.1	15.8	21.1	42.1
EDC	N	0	2	4	3	5	2	2	4	5	3
	%	0	14.3	28.6	21.4	35.7	12.5	12.5	25.0	31.3	18.8
Pittsburgh	N	0	3	3	4	0	0	1	2	3	8
	%	0	30.0	40.0	40.0	0	0		14.3	21.4	57.1
SEDL	N	1	1	2	2	1	1	1	2	2	2
	%	14.3	14.3	28.6	28.6	14.3	12.5	12.5	25.0	25.0	25.0

TABLE MIII-7

FREQUENCY COUNTS OF RECEIPT OF TITLE I AND OTHER FEDERAL OR STATE FUNDS FOR FOLLOW THROUGH AND NON-FOLLOW THROUGH

Did School Receive Title I Funds?

		FT	NFT
Yes	N	210	172
	%	92.5	85.1
No	N	17	30
	%	7.5	14.9

Did School Receive Other Federal or State Funds?

		FT	NFT
Yes	N	171	143
	%	88.6	84.6
No	N	22	26
	%	11.4	15.4

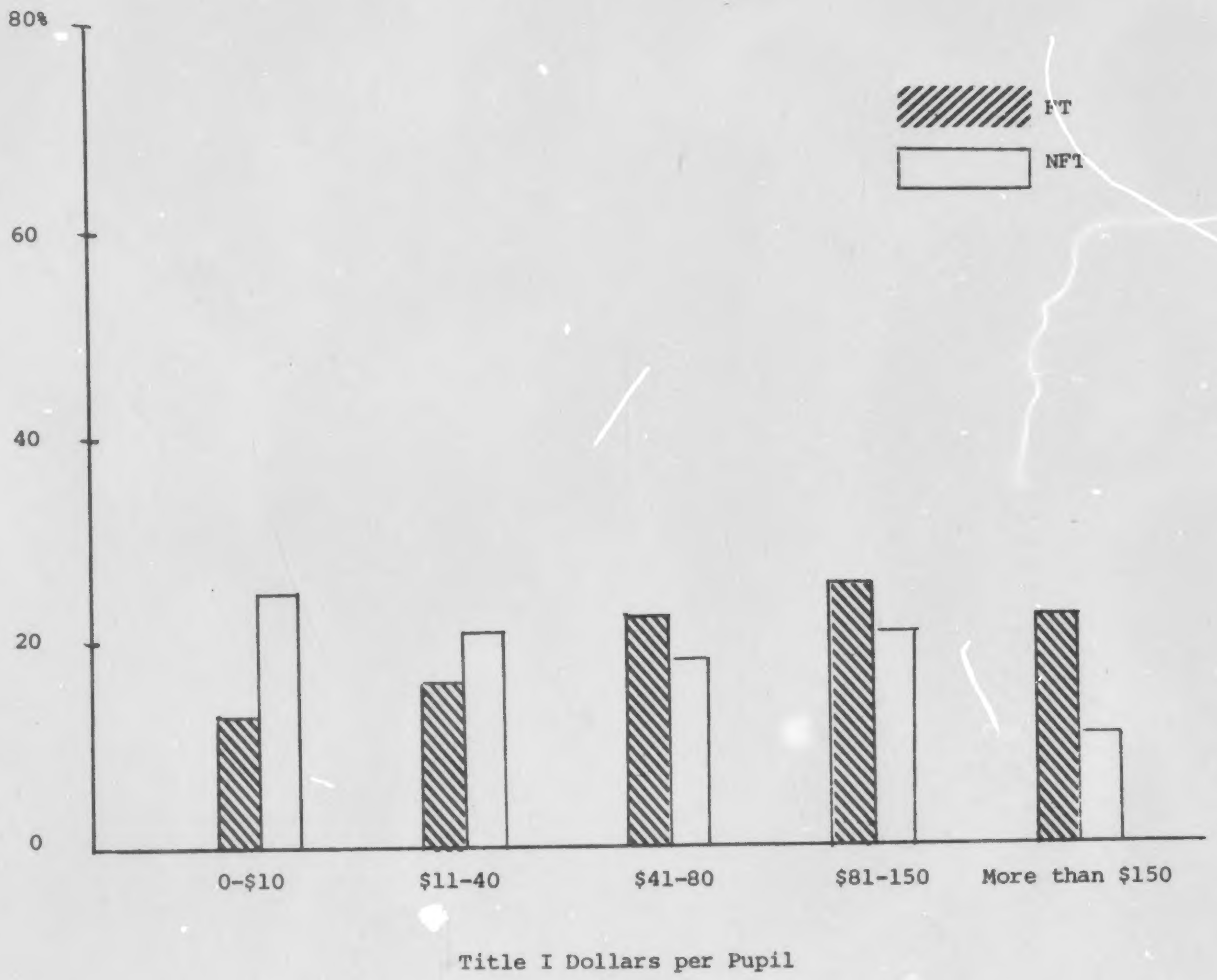


Figure MIII-5. Percentage of Title I Dollars per Pupil by FT/NFT

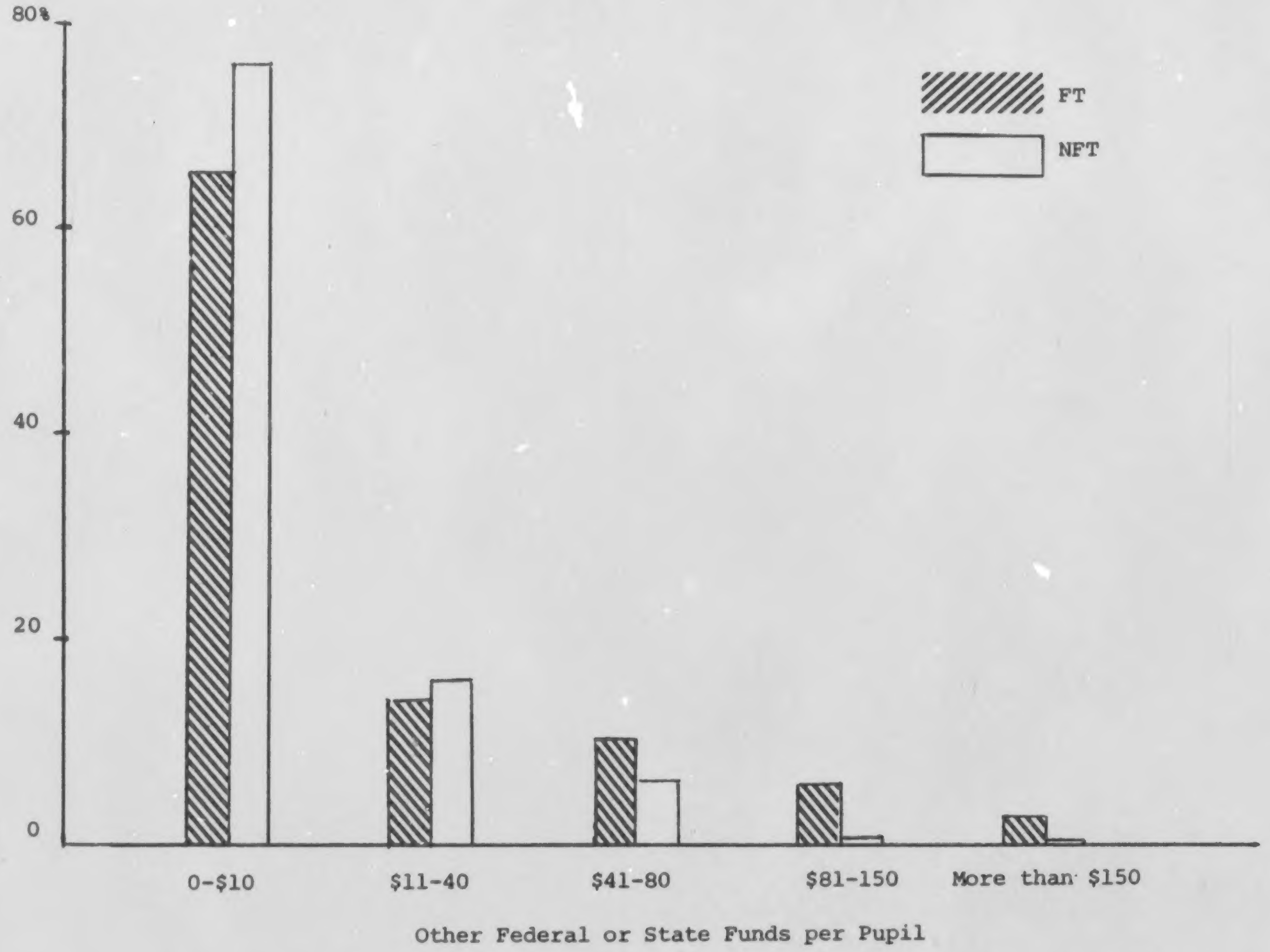


Figure MIII-6. Frequency Distributions of Other Federal or State Funds per Pupil for Follow Through and Non-Follow Through schools receive only small amounts of other Federal or State funds, indicating that few schools have programs that significantly add to Follow Through and/or Title I. Most of the schools receiving more than $80 per pupil of other Federal or State funds are FT schools.

With few exceptions, the overall FT/NFT pattern holds true across Sponsors. Only in the cases of Florida and Pittsburgh do a greater proportion of NFT schools than FT receive Title I. Florida has the smallest proportion of all FT schools receiving Title I, and Far West and Arizona have the smallest proportion of NFT schools. All FT schools for Oregon, Kansas, High/Scope and EDC receive Title I funds. For Bank Street, Florida, Kansas, and Pittsburgh, a greater proportion of NFT schools than FT receive other Federal or State funds. All FT schools for Arizona and High/Scope, and all NFT schools for Kansas receive these funds. All SEDL schools, both FT and NFT, receive both Title I and other Federal or State funds. (See Table 8).

Examining the amount of funds received by Sponsors, all Sponsors' FT schools, with the exception of Florida, receive more Title I dollars per pupil than their respective NFT schools (see Table 9). In the case of Florida, a slightly greater proportion of NFT schools than FT receive more than $80 per pupil (NFT=35.2 %; FT=28.1%). Far West schools receive the most Title I funds per pupil, with 53.6 percent of FT and 34.6 percent of NFT receiving more than $150 per pupil. Pittsburgh receives the least amount of Title I, with only 11.1 percent of FT schools and no NFT schools receiving more than $80 per pupil.

As shown in Table 10, all Sponsors' FT schools receive more Federal or State funds per pupil than their respective NFT schools except EDC and Pittsburgh where FT and NFT schools receive relatively low and equal amounts. Again, Far West receives the most funds, with 47 percent of FT schools and 13.4 percent of NFT schools receiving more than $80 per pupil. Kansas, High/Scope, EDC and Pittsburgh receive the smallest amount of other Federal or State funds for both FT and NFT schools.

TABLE MIII-8

FREQUENCY COUNTS OF RECEIPT OF TITLE I AND OTHER FEDERAL OR STATE FUNDS FOR EACH SPONSOR

		FT				NFT			
		Did School Receive Title I Funds?		Did School Receive Other Federal or State Funds?		Did School Receive Title I Funds?		Did School Receive Other Federal or State Funds?	
Sponsor		Yes	No	Yes	No	Yes	No	Yes	No
Far West	N	27	2	25	2	19	8	20	5
	%	93.1	6.9	92.6	7.4	70.4	29.6	80.0	20.0
Arizona	N	29	5	31	0	26	11	31	1
	%	85.3	14.7	100.0		70.3	29.7	96.9	3.1
Bank Street	N	31	1	16	7	20	3	11	6
	%	96.9	3.1	69.6	30.4	87.0	13.0	64.7	35.3
Oregon	N	25	0	18	1	25	3	20	6
	%	100.0		94.7	5.3	89.3	10.7	76.9	23.1
Kansas	N	23	0	16	3	16	1	12	0
	%	100.0		84.2	15.8	94.1	5.9	100.0	
High Scope	N	14	0	12	0	11	2	10	2
	%	100.0		100.0		84.6	15.4	83.3	16.7
Florida	N	25	7	24	3	19	1	14	1
	%	78.1	21.9	88.9	11.1	95.0	5.0	93.3	6.7
EDC	N	15	0	10	2	16	1	9	3
	%	100.0		83.3	16.7	94.1	5.9	75.0	25.0
Pittsburgh	N	10	2	8	4	13	0	9	2
	%	83.3	16.7	66.7	33.3	100.0		81.8	18.2
SEDL	N	11	0	11	0	7	0	7	0
	%	100.0		100.0		100.0		100.0	

TABLE MIII-9

FREQUENCY COUNTS OF TITLE I DOLLARS PER PUPIL FOR EACH SPONSOR

		PT Title I Dollars Per Pupil						NFT Title I Dollars Per Pupil					
Sponsor		0 to $10	$11 to $40	$41 to $80	$81 to $150	More than $150	N	0 to $10	$11 to $40	$41 to $80	$81 to $150	More than $150	N
Far West	N	5	1	1	6	15		10	0	1	6	9	
	%	17.9	3.6	3.6	21.4	53.6	28	38.5	0	0.6	23.1	34.6	26
Arizona	N	5	5	9	3	4		12	5	2	1	2	
	%	19.2	19.2	34.6	11.5	15.4	26	54.6	22.7	9.1	4.5	9.1	22
Bank Street	N	1	4	8	12	6		3	5	6	5	2	
	%	3.2	12.9	25.8	34.7	19.3	31	14.3	23.8	28.6	23.8	9.5	21
Oregon	N	1	3	3	7	2		8	5	3	7	0	
	%	6.3	18.8	18.8	43.8	12.5	16	34.8	21.7	13.0	30.4	0	23
Kansas	N	0	1	7	9	5		1	1	6	7	2	
	%	0	4.5	31.8	40.9	22.7	22	5.9	5.9	35.3	41.7	11.7	17
High Scope	N	0	5	2	3	2		2	6	4	1	0	
	%	0	41.7	16.7	25.0	16.7	12	15.4	46.2	30.8	7.7	0	13
Florida	N	11	6	6	4	5		5	3	3	3	3	
	%	34.3	18.7	18.7	12.5	15.6	32	29.4	17.6	17.6	17.6	17.6	17
EDC	N	0	1	5	4	3		1	2	4	5	1	
	%	0	7.7	38.4	30.7	23.1	13	7.7	15.4	30.7	38.4	7.7	13
Pittsburgh	N	2	1	5	1	0		0	4	1	0	0	
	%	22.2	11.1	55.5	11.1	0	9	0	80.0	20.0	0	0	5
SEDL	N	0	4	1	3	3		0	5	0	1	0	
	%	0	36.3	9.1	27.3	27.3	11	0	83.3	0	16.7	0	6

TABLE MIII-10

FREQUENCY COUNTS OF OTHER FEDERAL OR STATE DOLLARS PER PUPIL FOR EACH SPONSOR

		FT							NFT					
		Other Federal or State Dollars Per Pupil							Other Federal or State Dollars Per Pupil					
Sponsor		0 to $10	$11 to $40	$41 to $80	$81 to $150	More than $150	N		0 to $10	$11 to $40	$41 to $80	$81 to $150	More than $150	N
Far West	N	7	2	0	5	3		N	10	2	1	1	1	
	%	41.2	11.8	0	29.4	17.6	17	%	66.7	13.3	6.7	6.7	6.7	15
Arizona	N	15	1	3	0	0		N	20	0	0	0	0	
	%	78.9	5.2	15.8	0	0	19	%	100.0	0	0	0	0	20
Bank Street	N	8	5	7	2	0		N	9	5	0	0	0	
	%	36.3	22.7	31.8	9.1	0	22	%	64.3	35.7	0	0	0	14
Oregon	N	5	4	3	0	0		N	11	5	4	0	0	
	%	41.7	33.3	25.0	0	0	12	%	55.0	25.0	20.0	0	0	20
Kansas	N	17	0	0	1	0		N	10	0	2	0	0	
	%	94.4	0	0	5.6	0	18	%	83.3	0	16.7	0	0	12
High Scope	N	8	0	0	0	1		N	12	0	0	0	0	
	%	88.9	0	0	0	11.1	9	%	100.0	0	0	0	0	12
Florida	N	14	6	2	1	1		N	7	6	0	1	0	
	%	58.3	25.0	8.3	4.2	4.2	24	%	50.0	42.9	0	7.1	0	14
EDC	N	10	1	0	0	0		N	10	1	0	0	0	
	%	90.9	9.1	0	0	0	11	%	90.9	9.1	0	0	0	11
Pittsburgh	N	7	1	0	0	0		N	6	1	0	0	0	
	%	87.5	12.5	0	0	0	8	%	85.7	14.3	0	0	0	7
SEDL	N	7	1	1	1	0		N	3	2	1	0	0	
	%	70.0	10.0	10.0	10.0	0	10	%	50.0	35.3	16.7	0	0	6

3.3 School Services

As shown in Figure 7, overall, FT schools have more professional support services available than NFT schools. More FT schools provide more than three professionals per hundred children than NFT schools (FT=19.7%; NFT=10.4%). In the middle range of services (one and one-half to two professionals per hundred children), FT and NFT are almost equal (22% and 23.6% respectively). Finally, only 6.9% of the FT schools provide less than one professional per hundred children, whereas the NFT group has 25.8 percent providing this amount.

This FT/NFT pattern holds true for almost all Sponsors, (as seen in Table 11). For Oregon and Kansas, however, the pattern is reversed; FT schools have fewer professionals available than their respective NFT schools. Kansas, for example, has no FT schools providing more than two professionals per hundred children compared with 30.8 percent of NFT schools providing the same amount. For Oregon, 15 percent of FT schools provide more than two professionals per hundred FT children whereas 47.8 percent of the NFT schools provide this amount. Florida and EDC have only slightly more professionals overall for FT than NFT (both FT and NFT schools for Florida have relatively few), while Far West, Bank Street, and SEDL resemble the overall FT/NFT pattern. The greatest disparity between FT and NFT schools is found in Arizona, where FT schools have substantially more professionals available than NFT. Finally, for High/Scope and Pittsburgh, although FT schools have more professionals available than NFT schools, both groups have a substantial number of professional services.

Figures 8 through 13 display the frequency counts for each component of the professional services--medical, social, psychological, instructional, activities, and library for FT and NFT schools.

In medical services (see Figure 8), the majority of FT schools are represented in the mid-high range (0.3 or more professionals per hundred children) with no schools in the zero category. On the other hand, most of the NFT schools are in the low-mid range (0-0.2 professionals per hundred children). Table 12 illustrates that this difference is primarily a function of the greater availability of nurses in the FT group.

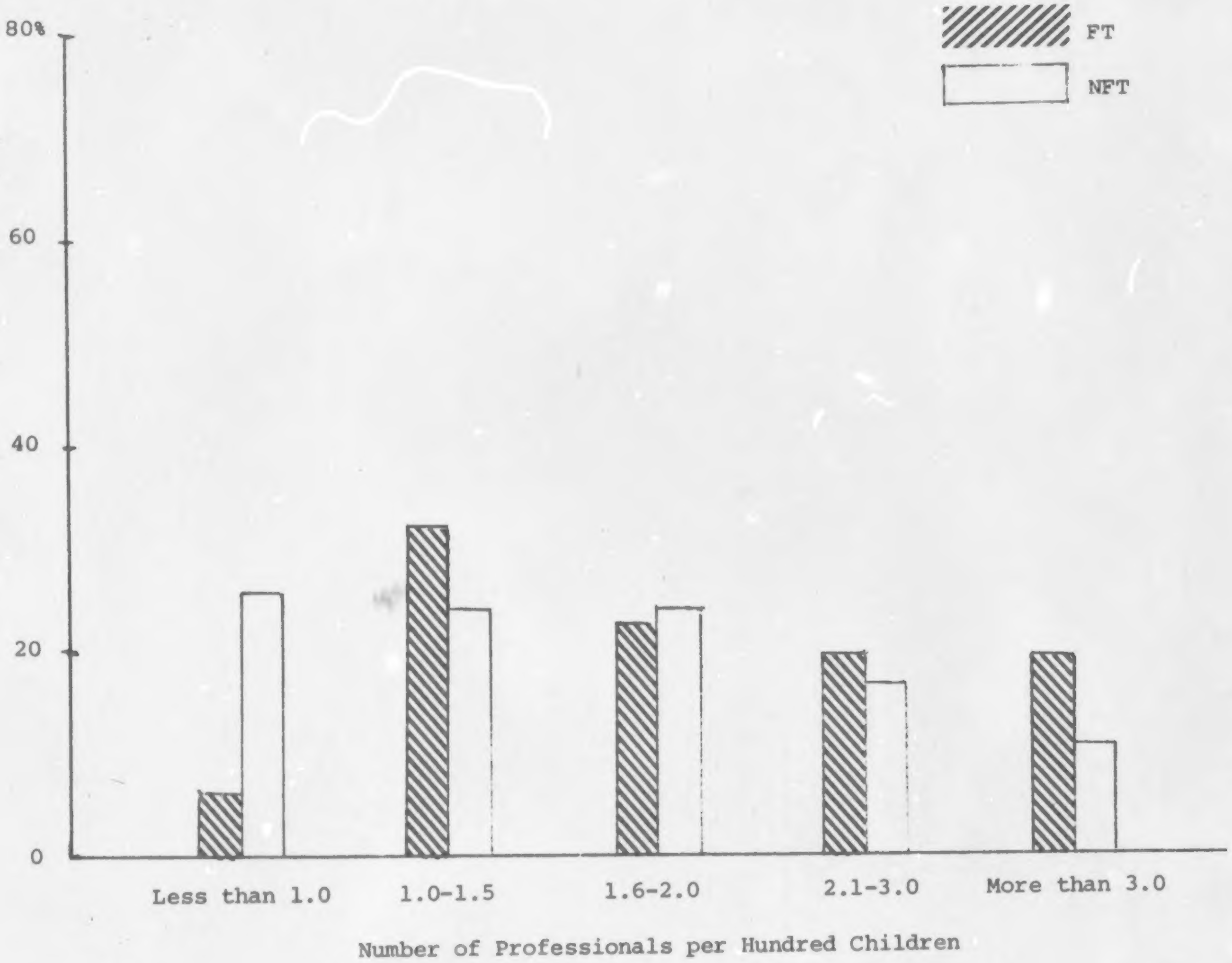


Figure MIII-7. Frequency Distributions of Professional Services for Follow Through and Non-Follow Through

TABLE MIII-11

FREQUENCY DISTRIBUTIONS OF PROFESSIONAL SERVICES FOR EACH SPONSOR

	FT						NFT					
	Number of Professionals Per Hundred Children						Number of Professionals Per Hundred Children					
Sponsor	Less than 1.0	1.0 - 1.4	1.6 - 2.0	2.2 - 3.0	More than 3.0	N	Less than 1.0	1.0 - 1.5	1.6 - 2.0	2.1 - 3.0	More than 3.0	N
Far West	1 3.3	8 26.7	8 26.7	7 23.3	6 20.0	30	9 33.3	6 22.2	8 29.6	3 11.1	1 3.7	27
Arizona	1 3.2	4 12.9	5 16.1	12 38.7	9 29.0	31	10 33.3	10 33.3	8 26.7	0 0	2 6.7	30
Bank Street	3 9.1	5 15.0	9 27.3	9 27.3	7 21.2	33	3 13.6	4 18.2	7 31.8	6 27.3	2 9.1	22
Oregon	1 5.0	10 50.0	6 30.0	2 10.0	1 5.0	20	3 13.0	6 26.1	3 13.0	6 26.1	5 21.7	23
Kansas	3 15.8	9 47.4	7 36.8	0 0	0 0	19	2 15.4	4 30.0	3 23.1	2 15.4	2 15.4	13
High Scope	0 0	3 21.4	0 0	3 21.4	8 57.1	14	2 16.7	1 8.3	4 33.3	2 16.7	3 25.0	12
Florida	4 13.3	16 53.3	4 13.3	3 10.0	3 10.0	30	8 42.1	5 26.3	2 10.5	3 15.8	1 5.3	19
EDC	2 11.1	6 33.3	6 33.3	3 16.7	1 5.6	18	4 26.7	5 33.3	4 26.7	2 13.3	0 0	15
Pittsburgh	0 0	3 25.0	0 0	1 8.3	8 66.7	12	1 7.1	3 21.4	3 21.4	5 35.7	2 14.3	14
SEDL	0 0	5 45.0	3 27.3	3 27.3	0 0	11	5 71.4	0 0	1 14.3	0 0	1 14.3	7

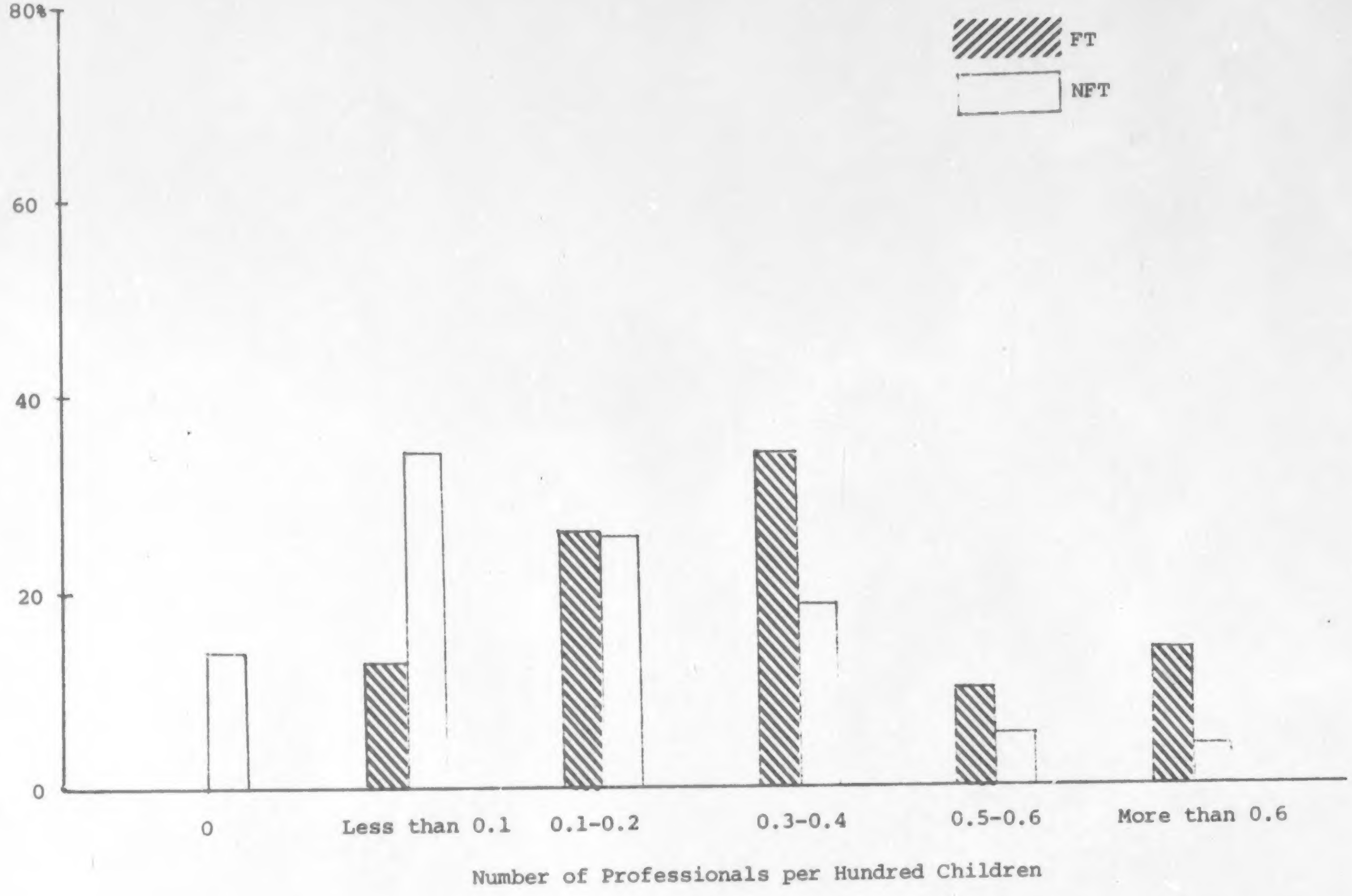


Figure MIII-8 Frequency Distributions of Medical Services for Follow Through and Non-Follow Through

TABLE MIII -12

FREQUENCY DISTRIBUTIONS OF MEDICAL SERVICES PERSONNEL FOR FOLLOW THROUGH AND NON-FOLLOW THROUGH

No. of Professionals Per Hundred Children		FT		NFT	
		Physician	Nurse	Physician	Nurse
less than 0.1	N	199	41	179	88
	%	91.3	18.8	98.4	48.3
0.1 - 0.4	N	15	138	3	58
	%	6.8	63.3	1.6	31.9
more than 0.4	N	4	39	0	36
	%	1.8	17.9	0	19.8

As Figure 9 illustrates, the majority of NFT schools have few social services available. Approximately 65 percent of the NFT schools provide less than 0.1 social service professionals per hundred children whereas 36.2 percent of FT schools provide this amount. In addition, 25.2 percent of FT and only 6.5 percent of NFT schools provide 0.5 or more professionals per hundred children. The FT advantage in social services can be attributed to the greater availability of social workers as shown in Table 13.

In the psychological component, FT schools have slightly more professionals available than NFT schools (see Figure 10). The NFT group has about half of its schools providing fewer than 0.1 professionals; the FT group has the majority of its schools providing more than this amount. The points at which the two groups differ markedly are in the zero category (FT = 12.8%; NFT = 28%) and in the 0.3-0.4 professionals category (FT = 27.5%; NFT = 17%). The overall FT/NFT difference can be attributed to more services in FT schools available from a psychologist or psychiatrist (see Table 14).

The FT group also exceeds the NFT group in the availability of instructional services (see Figure 11). Whereas 61 percent of the FT schools provide 0.7 or more professionals per hundred children, only 35.5 percent of the NFT schools provide this amount. The FT schools are fairly evenly distributed across all categories, whereas 31.9 percent of the NFT schools fall in the 0.1-0.4 category.

Table 15 illustrates the difference in the composition of the instructional component. The FT schools have substantially more curriculum specialists available than the NFT schools. The curriculum specialist for FT schools is the major contributor to FT instructional services, whereas for most NFT schools the reading specialist is the major contributor. The other personnel in the instructional component account for only a small amount of the services available to either group.

The availability of personnel for music, art and physical education activities in FT and NFT schools is shown in Figure 12. NFT schools have slightly more "specialist" teachers available than FT schools; 34.1 percent of NFT schools versus 24.3 percent of the FT schools have 0.5 or more professionals available per hundred children. The activities are fairly evenly split between music-art and physical education for both groups.

FT and NFT have practically equal amounts of library service personnel

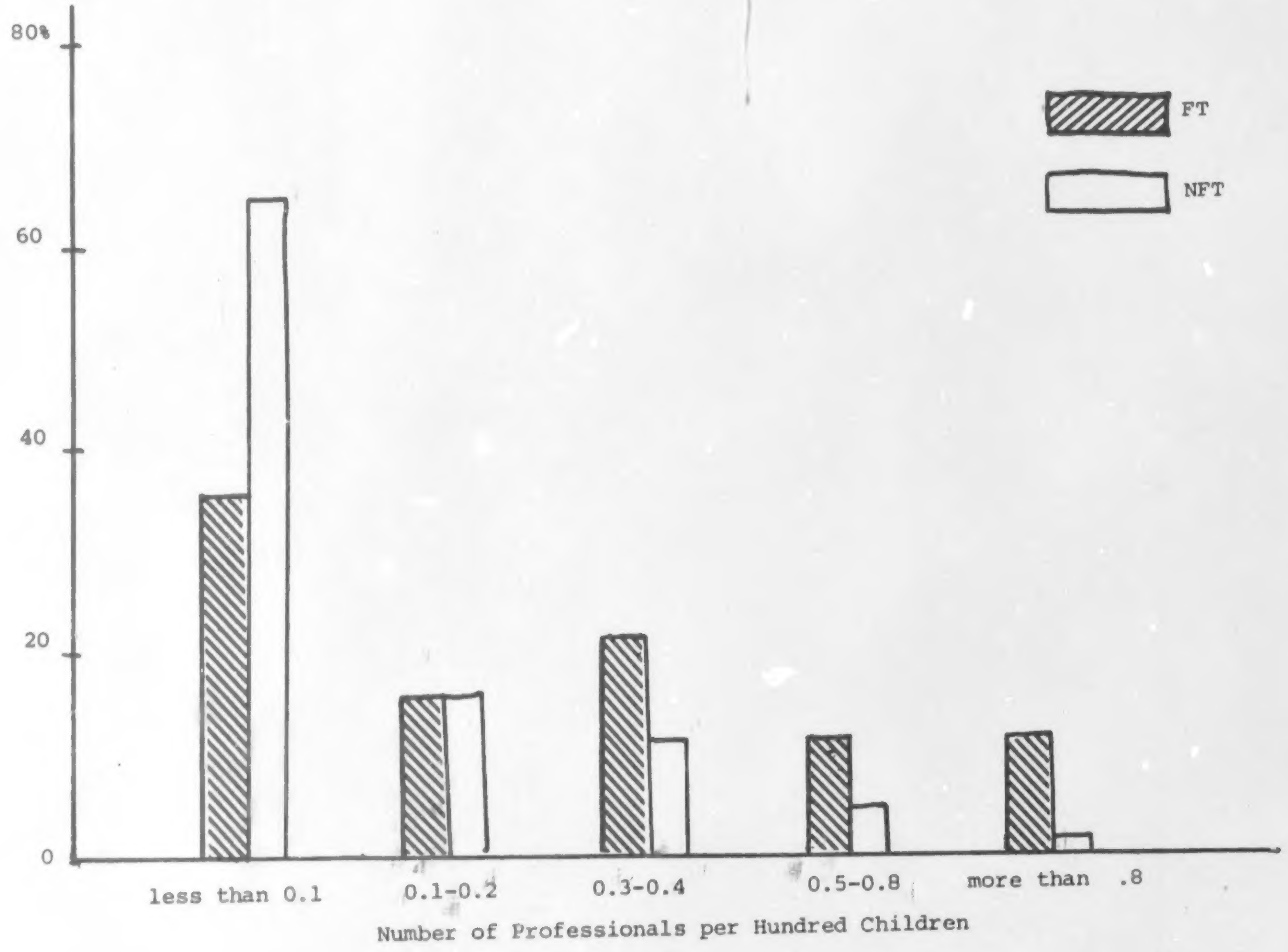


Figure MIII-9 Frequency Distributions of Social Services for Follow Through and Non-Follow Through

TABLE MI II-13

FREQUENCY DISTRIBUTIONS OF SOCIAL SERVICES PERSONNEL FOR FOLLOW THROUGH AND NON-FOLLOW THROUGH

No. of Professionals Per Hundred Children		FT		NFT	
		Social Worker	Child Welfare Worker	Social Worker	Child Welfare Worker
less than 0.1	N	87	164	146	148
	%	39.9	75.2	80.2	81.3
0.1 - 0.4	N	90	26	33	29
	%	41.3	11.9	18.2	15.9
more than 0.4	N	41	28	3	5
	%	18.8	12.8	1.6	2.7

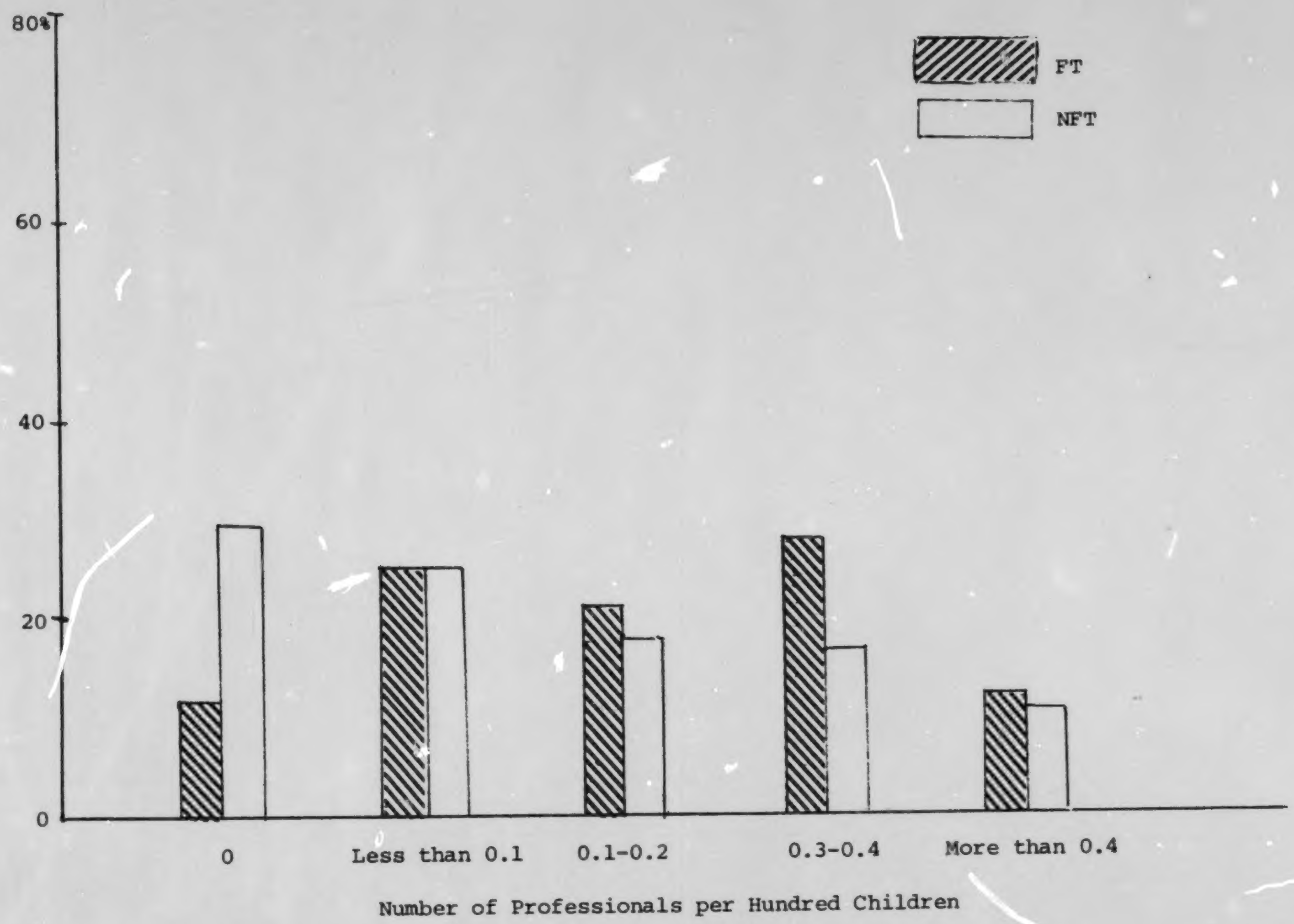


Figure MIII-10. Frequency Distributions of Psychological Services for Follow Through and Non-Follow Through

TABLE MIII-14

FREQUENCY DISTRIBUTIONS OF PSYCHOLOGICAL SERVICES PERSONNEL FOR FOLLOW THROUGH AND NON-FOLLOW THROUGH

No. of Professionals Per Hundred Children		FT		NFT	
		Psychologist/ Psychiatrist	Guidance Counselor	Psychologist/ Psychiatrist	Guidance Counselor
0-less than 0.1.	N	127	162	154	111
	%	58.3	74.3	84.6	61.0
0.1 - 0.4	N	91	56	28	57
	%	41.7	25.7	15.3	31.3
more than 0.4	N	0	0	0	14
	%	0	0	0	7.7

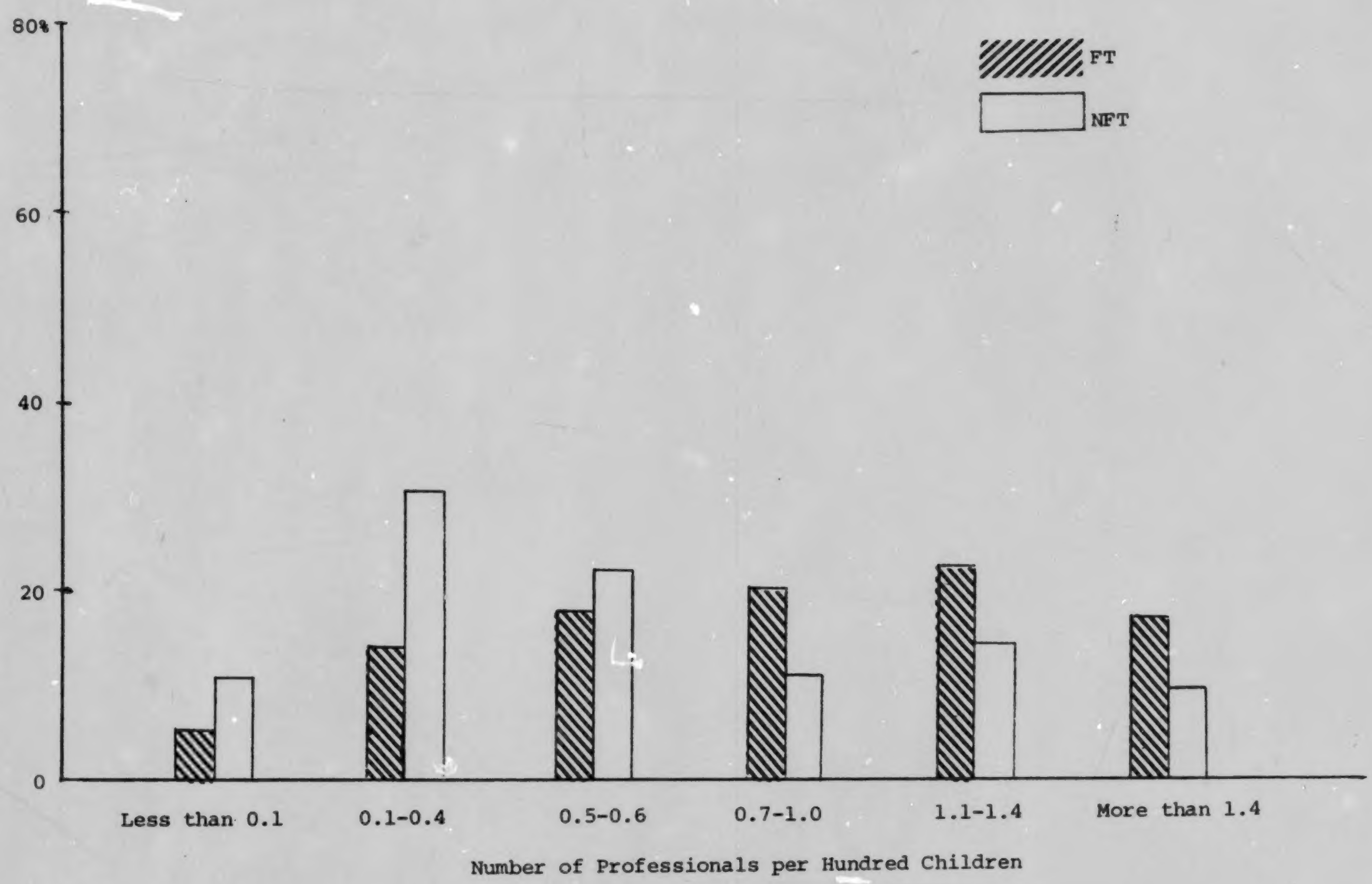


Figure MIII-11 Frequency Distributions of Instructional Services for Follow Through and Non-Follow Through

TABLE MIII-15

FREQUENCY DISTRIBUTIONS OF INSTRUCTION PERSONNEL

FT

No. of Professionals Per Hundred Children		Reading Specialist	Math Specialist	Curriculum Specialist	Speech Therapist	English-as-a-Second-Language Specialist
less than 0.5	N	194	214	101	211	208
	%	89.0	98.2	46.3	96.8	95.4
0.5 - 1.0	N	13	2	70	4	7
	%	6.0	0.9	32.1	1.8	3.2
more than 1.0	N	11	2	47	3	3
	%	5.0	0.9	21.6	1.4	1.4

N = 218

NFT

No. of Professionals Per Hundred Children		Reading Specialist	Math Specialist	Curriculum Specialist	Speech Therapist	English-as-a-Second-Language Specialist
less than 0.5	N	135	170	163	177	176
	%	74.2	93.5	89.6	97.4	96.8
0.5 - 1.0	N	15	2	7	1	2
	%	17.6	5.5	6.6	2.1	2.1
more than 1.0	N	15	2	7	1	2
	%	8.2	1.0	3.8	0.5	1.0

N = 182

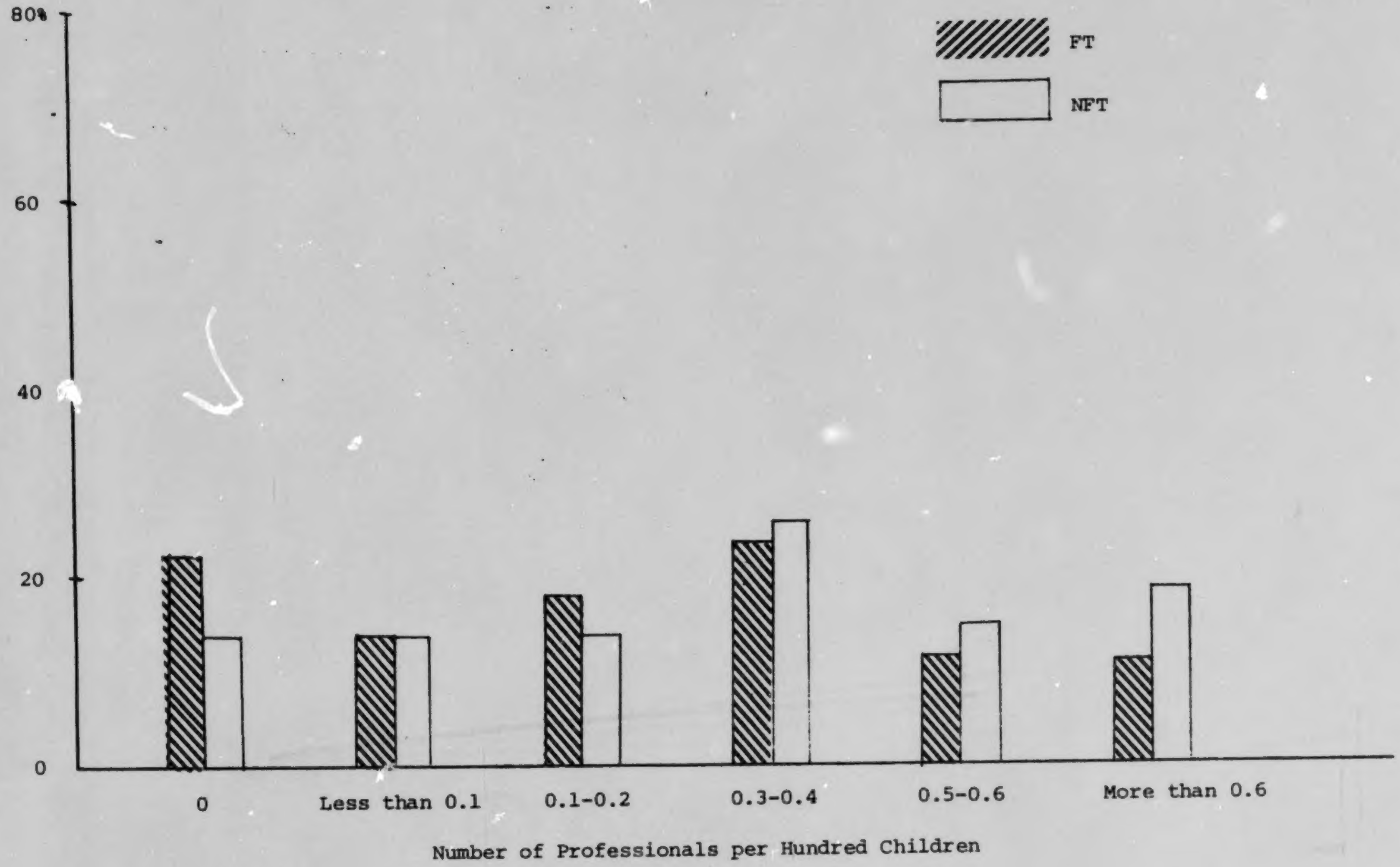


Figure MIII-12 Frequency Distributions of Activities for Follow Through and Non-Follow Through available. As shown in Figure 13, FT and NFT have equal proportions over the low, middle and high ranges of the distribution. Both groups have about half of the schools represented in the 0.1-0.4 professionals category. This component consists mostly of services of a librarian rather than an audiovisual specialist for both FT and NFT schools.

Figure 14 illustrates the non-professional services provided by FT and NFT schools. Of the NFT schools, 62.6 percent have no non-professionals whereas only 17.1 percent of the FT schools have none. The majority of FT schools provide 0.2 or more non-professionals. The difference between the two groups can be attributed to substantially more community liaison persons available in FT schools as seen in Table 16.

To summarize, in medical, social, instructional, and non-professional components, FT schools have substantially more personnel available than NFT schools. FT schools also have slightly more personnel available to provide psychological services than do NFT schools. The NFT group has slightly more personnel available beyond the classroom teacher to provide curriculum-related activities. Finally, in library services, the FT and NFT distributions are equal.

By examing each service component as a percentage of professional services we find that both groups concentrate their resources in instructional services (see Table 17); 38.5 percent of professional services in FT schools can be attributed to instructional services, and 34 percent in NFT schools. As a second priority, FT provides medical and social service personnel (15 and 14.8%, respectively), whereas NFT provides extra-class instructors such as activities (22.1%) and library personnel (20.2%). Psychological services represent the smallest percentage of professional services for both groups (FT= 8.6%; NFT = 8.2%). In addition, a large percentage of total FT services (23.7%) are available from non-professionals, primarily community liaisons, whereas only 4.9 percent of NFT total services are represented by this component.

Almost all Sponsors follow the overall FT/NFT pattern of having more medical, social, psychological and instructional personnel available in FT schools than NFT. As mentioned earlier, High/Scope and Pittsburgh have the most services for both FT and NFT. Arizona also has relatively large amounts of services available in FT schools and relatively few in NFT. (See Table 18.)

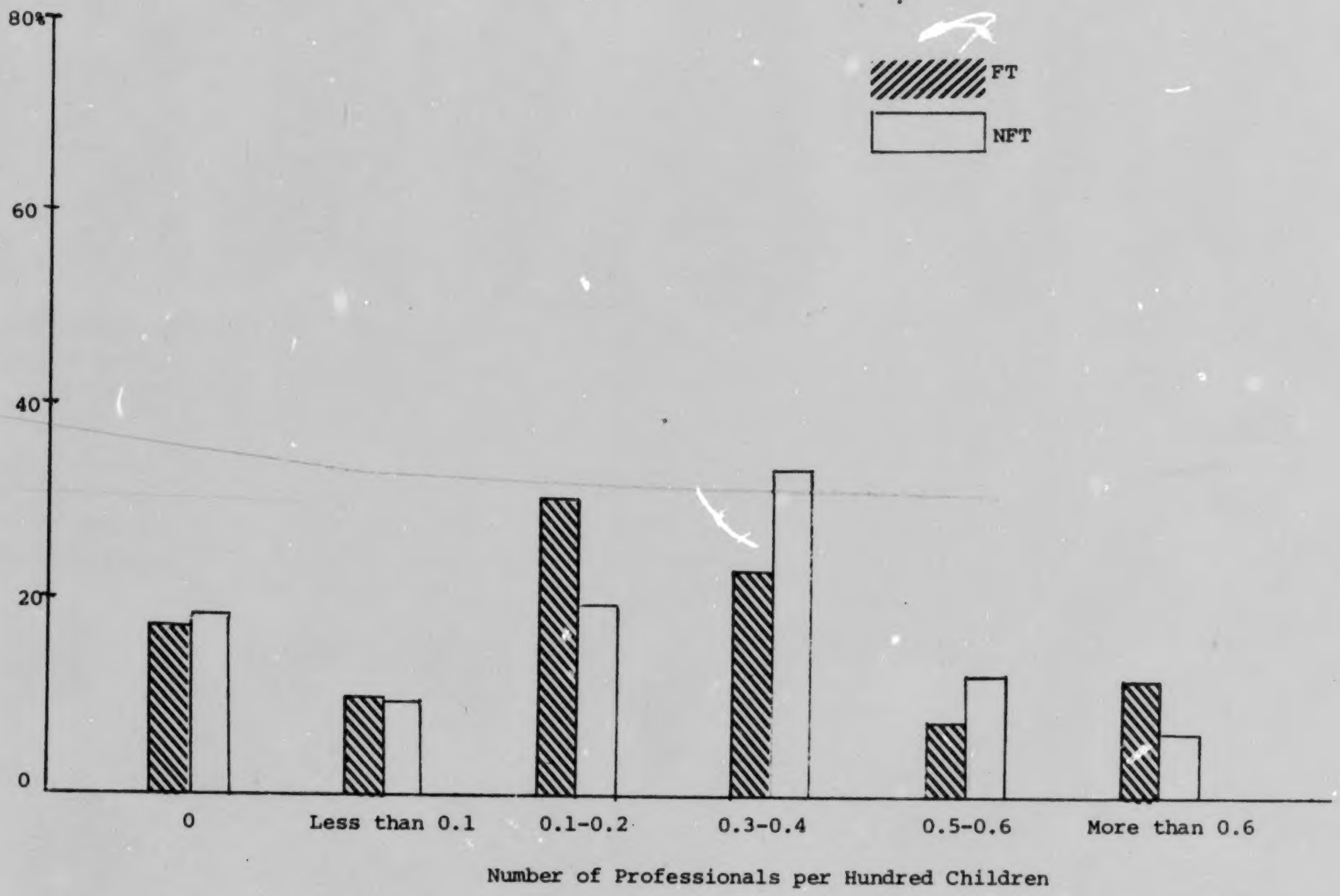


Figure MII-13 Frequency Distributions of Library Services for Follow Through and Non-Follow Through

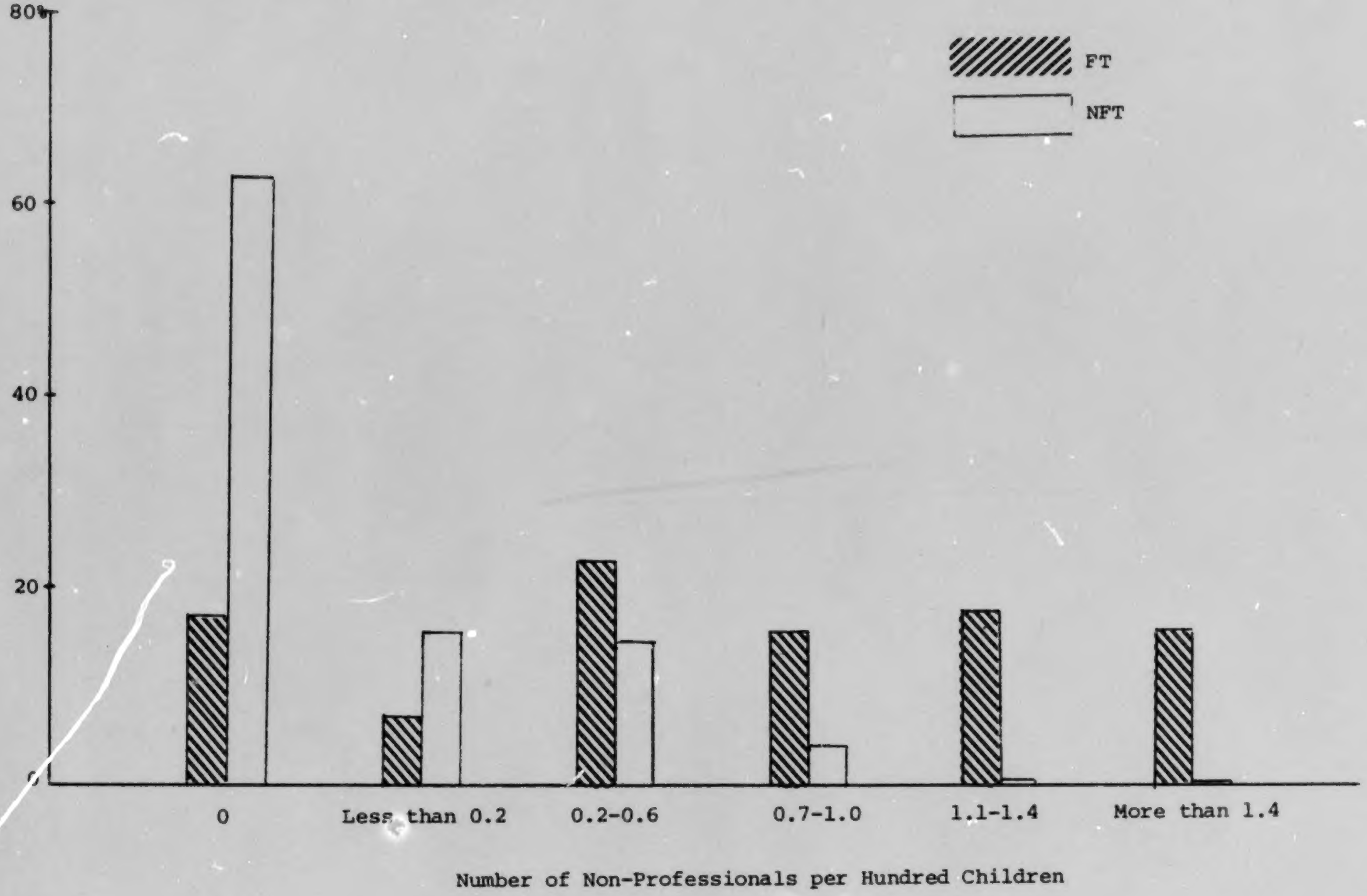


Figure MIII-14 Frequency Distributions for Non-Professional Services for Follow Through and Non-Follow Through

TABLE MIII-16

FREQUENCY DISTRIBUTIONS OF
NON-PROFESSIONAL SERVICES PERSONNEL FOR
FOLLOW THROUGH AND NON-FOLLOW THROUGH

No. of Professionals Per Hundred Children		FT		NFT	
		Community Liaison	Other Paraprofessionals	Community Liaison	Other Paraprofessionals
less than 0.2	N	69	169	150	172
	%	31.6	77.5	82.5	95.0
0.2 - 1.0	N	88	35	32	9
	%	40.4	16.1	17.5	5.0
more than 1.0	N	61	14	0	0
	%	28.0	6.4	0	0

TABLE MIII-17

SERVICE COMPONENTS AS PERCENTAGE OF TOTAL SERVICES FOR FT AND NFT

	Percent of Professional Services	
Service Component	FT	NFT
Medical	15.0	8.3
Social	14.8	7.0
Psychological	8.6	8.2
Instructional	38.5	34.0
Activities	12.0	22.1
Library	11.1	20.2
Non-Professional as Percent of Total Services	23.7	4.9

TABLE MIII-18

FREQUENCY DISTRIBUTIONS OF PROFESSIONAL SERVICES FOR EACH SPONSOR

FT

Number of Professionals Per Hundred Children	SPONSOR									
	Far West	Arizona	Bank Street	Oregon	Kansas	High Scope	Florida	EDC	Pitts-burgh	SEDL
Less than 1.0	1	1	3	1	3	0	4	2	0	0
	3.3	3.2	9.1	5.0	15.8	0	13.3	11.1	0	0
1.0-1.4	8	4	5	10	9	3	16	6	3	5
	26.7	12.9	15.1	50.0	47.4	21.4	53.3	33.3	25.0	45.4
1.6-2.0	8	5	9	6	7	0	4	6	0	3
	26.7	16.1	27.3	30.0	36.8	0	13.3	33.3	0	27.3
2.2-3.0	7	12	9	2	0	3	3	3	1	3
	23.3	38.7	27.3	10.0	0	21.4	10.0	16.7	8.3	27.3
More than 3.0	6	9	7	1	0	8	3	1	8	0
	20.0	29.0	21.2	5.0	0	57.1	10.0	5.6	66.7	0
N =	30	31	33	20	19	14	30	18	12	11

NFT

Number of Professionals Per Hundred Children	SPONSOR									
	Far West	Arizona	Bank Street	Oregon	Kansas	High Scope	Florida	EDC	Pitts-burgh	SEDL
Less than 1.0	9	10	3	3	2	2	8	4	1	5
	33.3	33.3	13.6	13.0	15.4	16.7	42.1	26.7	7.1	71.4
1.0-1.4	6	10	4	6	4	1	5	5	3	0
	22.2	33.3	18.2	26.1	30.	8.3	26.3	33.3	21.4	0
1.6-2.0	8	8	7	3	3	4	2	4	3	1
	29.6	26.7	31.8	13.0	23.1	33.3	10.5	26.7	21.4	14.3
2.2-3.0	3	0	6	6	2	2	3	2	5	0
	11.1	0	27.3	26.1	15.4	16.7	15.8	13.3	35.7	0
More than 3.0	1	2	2	5	2	3	1	0	2	1
	3.7	6.7	9.1	21.7	15.4	25.0	5.3	0	14.3	14.3
N =	27	30	22	23	13	12	19	15	14	7

Although there are more support services available in FT schools than NFT schools overall, there are also some interesting relationships between services and community variables. In the FT world, more social services are available in poor districts where a large percent of persons are below poverty level than in rich districts, whereas more medical services are available in rich districts than poor districts (r = .24 and -.25, respectively). In the NFT world, the largest correlations are those between non-professional services and three demographic variables related to city size. Although most NFT schools do not have non-professional services available, there is a tendency towards more non-professional services in large cities where density is high and elementary enrollments are high (r = .23, .35, and .23, respectively).

3.4 School District Characteristics

To describe the communities in which the schools are operating, we turn first to the occupational structure of the school district. As shown in Table 19, the majority of schools in this study are located in districts where about 14 percent of the working population are in professional occupations, 18 percent are operatives, and farm related occupations account for less than 5% of the persons in the district. Table 20 reveals some Sponsor to Sponsor variability. Far West schools are in communities with a large percentage of persons in farm occupations (17.3%) and a low percentage in professional occupations (10%). The large percentage of farm workers in SEDL communities can be attributed to one district that has 35 percent of its persons in farm labor occupations. High/Scope schools come from two types of districts. In one High/Scope district, for example, 20 percent of persons are in professional occupations. Pittsburgh also has two different types of districts. In one district where there are six schools, 15 percent of persons are in farm managerial occupations. Two other districts (with seven and six schools each) have 26 percent and 22 percent respectively in operative occupations. Relative to the other Sponsors, Pittsburgh has a high percentage of persons in both operative and farm related occupations. The remaining six Sponsors have similar occupation structures to the overall population with few exceptions. Oregon and Kansas have slightly more farm laborers than do the others; EDC has a slightly larger percentage of persons (24%) in operative occupations.

TABLE MIII-19

MEANS AND STANDARD DEVIATIONS OF PERCENT PERSONS IN CENSUS OCCUPATION CATEGORIES FOR SPONSORS OVERALL

Occupation Category	Overall	
	$\bar{X}$	S.D.
Professional	14.0	4.8
Managerial	7.3	2.1
Sales	6.7	1.8
Clerical	18.1	4.2
Craftsmen	12.5	2.2
Operatives	18.3	6.9
Laborers (not farm)	4.9	1.2
Service Workers	15.0	3.2
Farm Managers	1.2	2.8
Farm Laborers	1.9	5.6

N = 370

TABLE MIII-2C

MEANS AND STANDARD DEVIATIONS
OF PERCENT PERSONS IN CENSUS OCCUPATION CATEGORIES
FOR EACH SPONSOR

Occupation Category		SPONSOR									
		Far West	Arizona	Bank Street	Oregon	Kansas	High Scope	Florida	EDC	Pittsburgh	SEDL
Professional	$\overline{X}$	19.1	14.4	14.4	13.1	11.8	15.9	13.4	12.4	12.4	10.0
	S D.	8.6	3.2	5.3	3.3	1.7	4.2	1.7	4.5	1.4	2.4
Managerial	$\overline{X}$	7.8	8.4	6.8	6.6	5.3	8.0	8.3	6.1	8.3	6.3
	S.D.	2.1	1.9	1.7	2.4	1.4	1.9	1.5	2.0	2.0	1.9
Sales	$\overline{X}$	7.3	7.7	5.5	6.4	5.8	6.8	8.0	6.0	6.2	4.7
	S.D.	1.5	1.5	1.8	2.2	0.9	1.6	1.3	1.2	1.4	1.9
Clerical	$\overline{X}$	20.1	19.0	17.4	17.7	19.7	17.9	19.4	17.0	14.6	13.7
	S.D.	1.9	3.3	3.8	3.1	4.8	5.9	3.1	4.7	2.5	7.1
Craftsmen	$\overline{X}$	11.7	13.1	12.4	12.4	11.5	11.8	13.0	13.6	13.2	11.6
	S.D.	3.1	1.6	2.3	1.8	1.3	1.9	1.7	2.7	2.0	3.7
Operatives	$\overline{X}$	14.0	17.8	19.3	20.0	20.7	13.4	16.5	24.0	21.3	17.3
	S.D.	5.3	7.7	7.3	6.6	4.5	4.1	4.3	9.1	4.7	3.7
Laborers (not farm)	$\overline{X}$	4.7	4.5	5.2	4.9	5.1	4.2	5.0	4.5	5.3	5.6
	S.D.	1.0	1.0	1.3	1.5	1.1	0.4	0.6	0.7	1.9	2.0
Service Workers	$\overline{X}$	14.7	14.2	17.5	15.7	16.2	15.1	15.2	13.2	13.1	13.6
	S.D.	1.2	2.7	4.5	4.2	2.3	1.3	2.0	2.7	1.6	3.6
Farm Managers	$\overline{X}$	0.1	0.5	0.5	1.1	1.5	1.4	0.3	2.0	4.3	3.7
	S.D.	0.1	1.3	0.7	2.4	3.9	1.7	0.4	3.6	6.5	3.1
Farm Laborers	$\overline{X}$	0.5	0.4	0.8	2.1	2.3	5.4	0.8	1.2	1.3	13.6
	S.D.	0.6	0.5	0.9	5.8	5.4	8.5	1.1	1.9	1.3	18.1
N = 370		39	63	48	47	30	27	47	31	24	14

The percentage of females in the labor force for the overall sample is 42.3%. The Sponsors are all very similar on this variable except for Pittsburgh and SEDL. These Sponsors, which also have large percentages of persons in farm-related occupations, have slightly fewer females in the labor force (as seen in Table 21).

There are only slight differences across Sponsors in the average income of the districts. Far West and Arizona districts have the highest average incomes ($10,840 and $10,384, respectively) and SEDL and Pittsburgh have the lowest ($8,309 and $8,789, respectively). The other Sponsors have an average family income of approximately $9,500 for the districts in which their schools are located (see Table 22).

Turning to mean education level of the districts, Table 23, Far West again has the highest average. The average number of years of education for Far West is 11.4, with one district having an average of 13.5 years. SEDL has the lowest average with only 8.5 years of education (one SEDL district has 6.9 years, the lowest for any Sponsor).

Table 24 presents district mobility, defined as the percentage of persons in the district who have always lived in the same house. Far West, Arizona, High/Scope and Florida have the smallest percentage of persons who have always lived in the same house (less than 5%). EDC and SEDL have the largest percentage (about 9%). The other Sponsors have means similar to the overall population (about 6-7%).

Far West and Arizona have the smallest percentage (10%) of persons below poverty level in the districts in which their schools are located as seen in Table 25. High/Scope and SEDL have the largest percentage of persons below poverty level (20.7% and 21.9%, respectively), indicating large low-income populations. One High/Scope district has 48 percentage of persons below poverty level and one SEDL district has 36 percent, the two largest percentages for all districts. High/Scope and Pittsburgh appear to have some discrepancy in the two income indices (mean family income and percentage of persons below poverty level), but these are also the Sponsors whose occupational structures indicated two different types of districts.

Pittsburgh has the lowest percentage (9.4%) of families with a female head of household and children under 18; Kansas has the largest percentage

TABLE MIII-21

MEANS AND STANDARD DEVIATIONS OF PERCENT FEMALES IN LABOR FORCE FOR EACH SPONSOR

Sponsor		Percent Females in Labor Force
Far West	$\bar{X}$	42.7
	S.D.	4.2
Arizona	$\bar{X}$	44.5
	S.D.	5.5
Bank Street	$\bar{X}$	43.0
	S.D.	3.6
Oregon	$\bar{X}$	40.8
	S.D.	5.2
Kansas	$\bar{X}$	43.1
	S.D.	6.4
High/Scope	$\bar{X}$	40.1
	S.D.	5.2
Florida	$\bar{X}$	43.8
	S.D.	2.9
EDC	$\bar{X}$	43.7
	S.D.	3.5
Pittsburgh	$\bar{X}$	36.8
	S.D.	3.5
SEDL	$\bar{X}$	37.2
	S.D.	7.8
Overall	$\bar{X}$	42.3
	S.D.	5.2

TABLE MIII-22

MEANS AND STANDARD DEVIATIONS OF AVERAGE FAMILY INCOME FOR EACH SPONSOR

Sponsor		Average Family Income
Far West	$\overline{X}$	10,839.69
	S.D.	810.42
Arizona	$\overline{X}$	10,383.58
	S.D.	1,227.92
Bank Street	$\overline{X}$	9,528.88
	S.D.	1,703.24
Oregon	$\overline{X}$	9,553.55
	S.D.	2,012.48
Kansas	$\overline{X}$	9,552.98
	S.D.	1,306.51
High/Scope	$\overline{X}$	9,271.74
	S.D.	2,254.48
Florida	$\overline{X}$	9,828.02
	S.D.	982.38
EDC	$\overline{X}$	9,640.86
	S.D.	1,558.58
Pittsburgh	$\overline{X}$	8,789.34
	S.D.	1,575.31
SEDL	$\overline{X}$	8,308.52
	S.D.	1,820.85
Overall	$\overline{X}$	9,752.13
	S.D.	1,625.66

TABLE MIII-23

MEANS AND STANDARD DEVIATIONS OF AVERAGE YEARS OF EDUCATION FOR EACH SPONSOR

Sponsor		Average Years of Education
Far West	$\overline{X}$	11.40
	S.D.	1.20
Arizona	$\overline{X}$	10.73
	S.D.	1.05
Bank Street	$\overline{X}$	10.11
	S.D.	1.25
Oregon	$\overline{X}$	9.92
	S.D.	0.98
Kansas	$\overline{X}$	9.58
	S.D.	0.71
High/Scope	$\overline{X}$	10.30
	S.D.	1.53
Florida	$\overline{X}$	10.39
	S.D.	0.39
EDC	$\overline{X}$	9.95
	S.D.	0.99
Pittsburgh	$\overline{X}$	10.48
	S.D.	0.46
SEDL	$\overline{X}$	8.55
	S.D.	1.29
Overall	$\overline{X}$	10.29
	S.D.	1.17

TABLE MIII-24

MEANS AND STANDARD DEVIATIONS OF PERCENT PERSONS ALWAYS LIVED IN SAME HOUSE FOR EACH SPONSOR

Sponsor		Percent Persons Always Lived in Same House
Far West	$\bar{X}$	4.7
	S.D.	1.3
Arizona	$\bar{X}$	4.8
	S.D.	1.4
Bank Street	$\bar{X}$	7.6
	S.D.	2.4
Oregon	$\bar{X}$	7.9
	S.D.	3.9
Kansas	$\bar{X}$	6.3
	S.D.	1.6
High/Scope	$\bar{X}$	4.8
	S.D.	1.9
Florida	$\bar{X}$	5.0
	S.D.	1.4
EDC	$\bar{X}$	8.7
	S.D.	3.1
Pittsburgh	$\bar{X}$	7.1
	S.D.	1.9
SEDL	$\bar{X}$	9.0
	S.D.	5.1
Overall	$\bar{X}$	6.3
	S.D.	2.9

TABLE MIII-25

MEANS AND STANDARD DEVIATIONS OF PERCENT PERSONS BELOW POVERTY LEVEL FOR EACH SPONSOR

Sponsor	Percent Persons Below Poverty Level	
Far West	$\overline{X}$	10.5
	S.D.	1.2
Arizona	$\overline{X}$	10.8
	S.D.	4.3
Bank Street	$\overline{X}$	16.0
	S.D.	10.8
Oregon	$\overline{X}$	17.8
	S.D.	10.3
Kansas	$\overline{X}$	16.2
	S.D.	8.5
High/Scope	$\overline{X}$	20.7
	S.D.	17.1
Florida	$\overline{X}$	14.2
	S.D.	3.0
EDC	$\overline{X}$	12.0
	S.D.	6.1
Pittsburgh	$\overline{X}$	13.4
	S.D.	6.1
SEDL	$\overline{X}$	21.9
	S.D.	10.6
Overall	$\overline{X}$	14.6
	S.D.	8.9

(20.5%), SEDL, EDC, and Arizona are slightly below the overall mean with about 13 percent each; Far West, Bank Street, and Oregon are slightly above the overall mean with 17-18 percent each (see Table 26).

The two language indices displayed in Table 27 show some interesting Sponsor differences. Oregon and SEDL have the largest Spanish-speaking and the largest non-English speaking populations. For Oregon, this can be attributed to two districts with substantial Spanish-speaking populations (54.1% and 75.5%). In the case of SEDL it can be attributed to one district where 43.8% of persons are Spanish-speaking and a second district where 84% are non-English speaking. The other Sponsors show only slight variation in the language spoken in the district; Florida has the smallest non-English speaking population (less than 10%).

In sum, there are many Sponsor to Sponsor differences in the districts where the schools are located. Far West schools are located in communities where a large percentage of persons are in professional occupations, income and education levels are quite high, and the population is fairly mobile, compared with other Sponsor districts. Far West also has a relatively large percentage (18.1%) of families with a female head of household and children under 18. Arizona and Bank Street are similar to each other and the overall Sponsor mean in occupational structure, education level, and native language. Arizona districts have a higher income level, and more mobility than Bank Street districts; however, Bank Street districts have slightly more persons below poverty level than the average and more families with female heads of household, whereas Arizona has slightly fewer. Oregon and Kansas districts also are quite similar in all characteristics except for Oregon's large Spanish-speaking population and Kansas' large percentage of families with female heads of household. High/Scope is found in two very different types of districts. One type has a relatively large percentage of professionals; the other has a large number of farm laborers and a very high percentage of persons below poverty level. Florida most closely resembles the overall mean, with most persons in operative, clerical, and service occupations; an average income of $9,828; 14.2 percent of persons below poverty level; and an average of 10.4 years of education. This Sponsor differs markedly from the overall mean only in that it has the smallest percentage of persons in operative oc-

TABLE MIII-26

MEANS AND STANDARD DEVIATIONS OF PERCENT FAMILIES WITH CHILDREN UNDER 18 WITH FEMALE HEADS OF HOUSEHOLD FOR EACH SPONSOR

Sponsor	Percent Families with Children Under 18 with Female Heads of Household	
Far West	$\overline{X}$	18.1
	S.D.	4.1
Arizona	$\overline{X}$	13.8
	S.D.	5.9
Bank Street	$\overline{X}$	17.5
	S.D.	6.4
Oregon	$\overline{X}$	17.3
	S.D.	6.6
Kansas	$\overline{X}$	20.5
	S.D.	4.6
High/Scope	$\overline{X}$	15.2
	S.D.	3.9
Florida	$\overline{X}$	16.5
	S.D.	4.0
EDC	$\overline{X}$	13.4
	S.D.	6.5
Pittsburgh	$\overline{X}$	9.4
	S.D.	3.0
SEDL	$\overline{X}$	13.2
	S.D.	5.1
Overall	$\overline{X}$	15.8
	S.D.	5.9

TABLE MIII-27

MEANS AND STANDARD DEVIATIONS OF PERCENT SPANISH SPEAKING AND PERCENT NON-ENGLISH SPEAKING PERSONS FOR EACH SPONSOR

Sponsor		Percent Spanish Speaking	Percent Non-English Speaking
Far West	$\bar{X}$	4.3	21.9
	S.D.	4.6	5.4
Arizona	$\bar{X}$	2.2	10.4
	S.D.	2.8	8.4
Bank Street	$\bar{X}$	1.4	16.8
	S.D.	2.3	15.4
Oregon	$\bar{X}$	15.3	28.7
	S.D.	26.8	23.8
Kansas	$\bar{X}$	3.2	17.2
	S.D.	4.4	14.5
High/Scope	$\bar{X}$	5.3	15.7
	S.D.	5.4	14.2
Florida	$\bar{X}$	2.5	8.9
	S.D.	3.9	7.3
EDC	$\bar{X}$	2.1	20.6
	S.D.	4.2	14.8
Pittsburgh	$\bar{X}$	0.2	13.1
	S.D.	0.2	10.8
SEDL	$\bar{X}$	16.4	50.0
	S.D.	22.4	26.8
Overall	$\bar{X}$	4.7	18.0
	S.D.	1.2	16.8

cupations, relatively low education levels, a large percentage of persons who have always lived in the same house. Pittsburgh districts appear to be about average in income, education level and mobility, with relatively few women in the labor force or female heads of household and a small percentage of non-English speaking persons. SEDL appears to be in districts which contain a large percentage of person disadvantaged by occupation, income and language. Many persons in SEDL communities are in farm occupations, with low incomes and education levels: relatively few women work or are heads of household, and a large percentage of persons have always lived in the same house.

4.0 DISCUSSION

Two findings in this study are worthy of special note. First, we found little difference between FT and NFT in the proportion of schools receiving Title I and/or Federal or State funds. This finding suggests first of all that a reasonably close match was attained between the schools served by FT and NFT. In both cases, the schools serve low-income children eligible for Federal funding. This finding also poses some serious problems for the interpretation of FT/NFT contrasts. FT is not being compared to a "regular" school program, but rather to supplementary programs over and above the basic level of effort provided by the local educational agency. The quality of such programs varies widely, suggesting that additional information on the comparison programs would greatly enhance our understanding of pupil results.

The second finding of note is the FT/NFT difference in the method of delivery of support services. The most available professionals in the FT world are the curriculum specialists, community liaisons, and social workers, whereas in the NFT world specialist teachers are more common. In the FT model of delivery, the teacher or parent provides services directly to the children. The teacher, in turn, receives support services from the curriculum specialist and the parent from a social worker and/or a community liaison person. Hence, FT support services do not alter the classroom environment, but rather aid parents and teachers who are in direct contact with the children. The NFT model is more dependent on specialists and/or activities that service the children directly (i.e., supplementing or replacing regular classroom teaching).

The FT method is designed to be less disruptive to the pupil-teacher relationship than the NFT. Furthermore, in the FT pattern of delivery, those children who may be in need of special services are not obviously segregated from the rest of the class. The FT method also extends beyond the school to the home. These two findings, as well as other differences in the schools and communities in which FT Sponsors operate, must be considered in interpreting pupil results. The impact of these differences on pupil outcomes cannot be assessed until further studies merge school/community and pupil data.

MONOGRAPH IV

FOLLOWING THROUGH ON HEAD START VARIATIONS

MONOGRAPH IV

FOLLOWING THROUGH ON HEAD START PLANNED VARIATIONS

1.0 THE ORIGINS OF HEAD START PLANNED VARIATIONS

In the context of a major report on early effects of Follow Through, we need say little more about the genesis and motivation of Head Start Planned Variations (HSPV) than to point out that it arose from a very logical union of the philosophies of Follow Through and its parent project, Head Start; (Smith (1973) and Weisberg (1973) provide more historical details for the interested reader). In 1969, the concept of sponsorship was extended experimentally back from the early primary grades into Head Start, in order to seek answers to several questions about the role of preschool compensatory experience in later school performance:

- Does continuous experience with one Sponsor from Head Start through Follow Through enable children to maintain and build upon Head Start gains better than if their Head Start experience were unsponsored? better than if their Head Start and their Follow Through experiences had different Sponsors?
- Ignoring Head Start gains, does continuous sponsorship make children more successful in the early primary grades than their peers who had no Head Start, unsponsored Head Start, or Head Start under another Sponsor?
- Are some Sponsor models more effective overall than others? Is each model particularly effective for some kinds of children or in some kinds of locations?

In order to answer these questions, a battery of tests was administered to children in HSPV programs, to some comparison children in unsponsored Head Start (NPV) programs, and to a group of "Control" children who had no preschool experience. The sites for this HSPV testing were chosen from among existing Head Start sites in areas served by existing Follow Through programs. To qualify for participation, the Head Start programs had to agree to operate under the sponsorship of the local Follow Through Sponsor.

The studies of Smith (1973) and Weisberg (1973), already cited, examine the preschool effects of HSPV for the cohorts of children in Head Start in 1970-71 and in 1971-72, respectively. They were able to discern meaningful effects of the Head Start experience and some patterns of differential Sponsor performance, but they found no evidence that Sponsorship as such produced greater progress during the Head Start year. Without information about the later performance of the HSPV children in Follow Through, of course, Smith and Weisberg were unable to address the principal longitudinal issues of HSPV.

In the present report, for the first time, we have available a body of data on HSPV children traced into Follow Through Cohorts III and IV entering kindergarten, respectively. To the extent that enough data on the right groups have survived the tracing procedure, we have here the first opportunities to address the questions which motivated HSPV in the first place.

2.0 THE HSPV/FT STUDIES: INTRODUCTION

As we shall demonstrate, the surviving data provide the means to accomplish only a small proportion of the original purposes of the HSPV/FT longitudinal experiment. In particular, the quantity of pre-school test data on comparison children is far too small to support the kinds of conclusions we had hoped to draw concerning the gains of "treated" children relative to those of their "untreated" peers. We therefore present in the following sections one approximation to such a comparison, followed by such evidence as we can bring to bear on the secondary questions of differential sponsor performance. In the remainder of this section, we describe the state of the data and the constraints that it imposes on analysis.

2.1 Participating Groups

The 1272 children who took part in HSPV and then were later enrolled in one or two years of Follow Through are identified with six sponsors in the two Follow Through cohort streams, as Figure MIV-1 shows.

Each of these traced children is associated with a single Sponsor. That is, we have practically no data on children who may have participated in HSPV under one Sponsor and then in Follow Through under another Sponsor. We therefore have no way to compare the outcomes of continuity under a single Sponsor with those of experience under multiple sponsorship. The data we do have allow us to examine the expectation that continuous sponsorship enhances child performance, but they do not permit the comparison of continuously-sponsored children to other comparable children who lacked the HSPV experience.

Figure MIV-2 illustrates the distribution of the 416 traced children in Cohort III among the various paths that three successive stages of treatment/comparison distinctions define. Figure MIV-3 shows the analogous breakdown for Cohort IV*.

Given the substantial and important variance among Sponsors and among sites, it is clear that such lopsided distributions could not provide

---

[*] Since the real interest in HSPV-FT lies in its longitudinality, we focus throughout the remainder of this report on the three-year Cohort III data rather than on the two-year experience of Cohort IV. After Figure MIV-3, no further Cohort IV data are presented.

Sponsor	Cohort III (entered kindergarten in 1971)	Cohort IV (entered kindergarten in 1972)
2. Far West	145	164
3. Arizona	74	141
5. Bank Street	0	150
8. Kansas	56	159
9. High/Scope	141	138
12. Pittsburgh	0	103
Multiple Sponsors		1
Cohort Totals	416	855

Figure MIV-1 Sponsor and Cohort Distribution of Traced Children

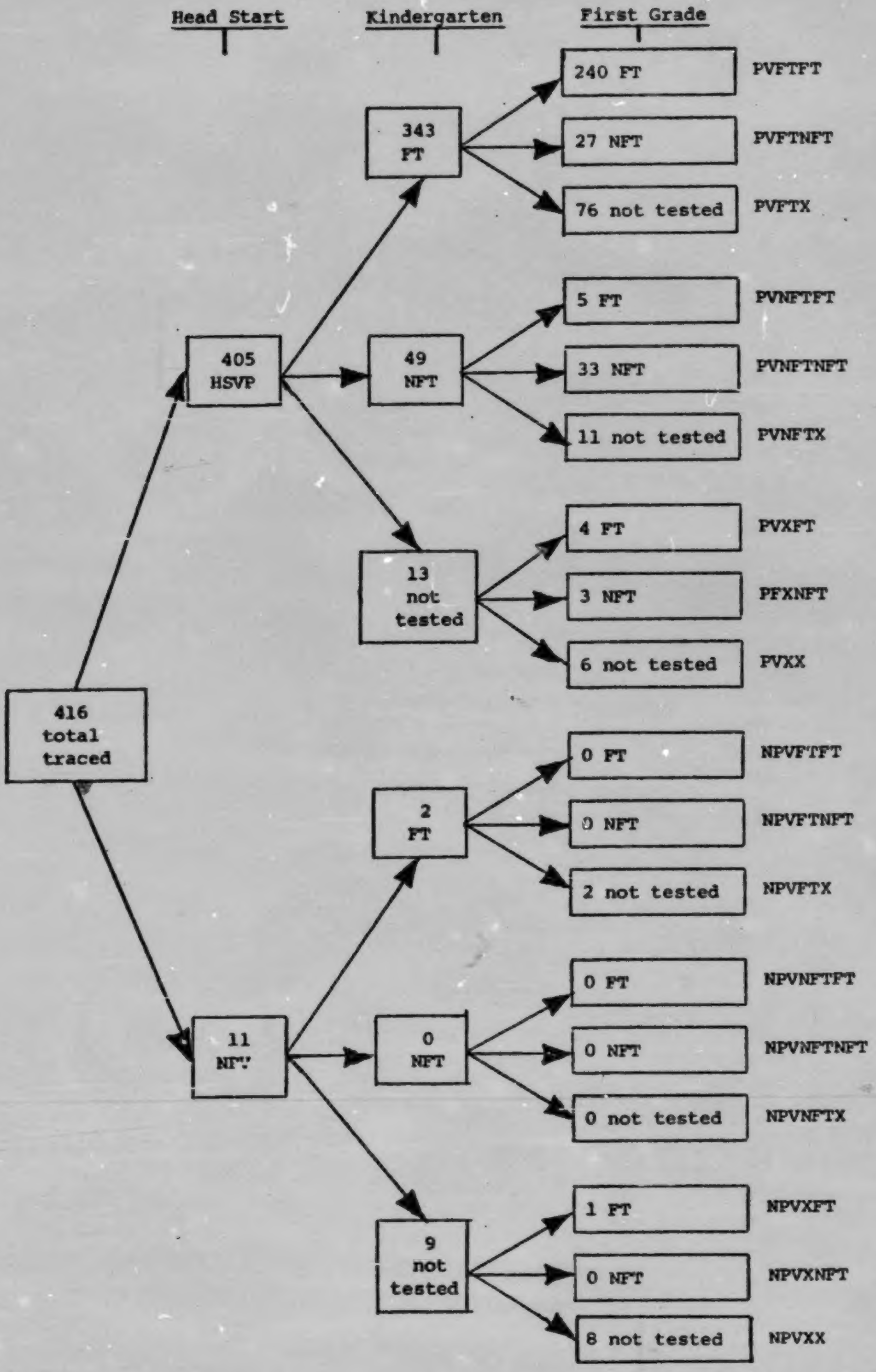

Head Start
Kindergarten
First Grade
416 total traced
405 HSVP
11 NPV
343 FT
49 NFT
13 not tested
2 FT
0 NFT
9 not tested
240 FT
27 NFT
76 not tested
5 FT
33 NFT
11 not tested
4 FT
3 NFT
6 not tested
0 FT
0 NFT
2 not tested
0 FT
0 NFT
0 not tested
1 FT
0 NFT
8 not tested
PVFTFT
PVFTNFT
PVFTX
PVNFTFT
PVNFTNFT
PVNFTX
PVXFT
PFXNFT
PVXX
NPVFTFT
NPVFTNFT
NPVFTX
NPVNFTFT
NPVNFTNFT
NPVNFTX
NPVXFT
NPVXNFT
NPVXX

Figure MIV-2. Treatment History Distribution of HSPV-FT Children in Cohort III.

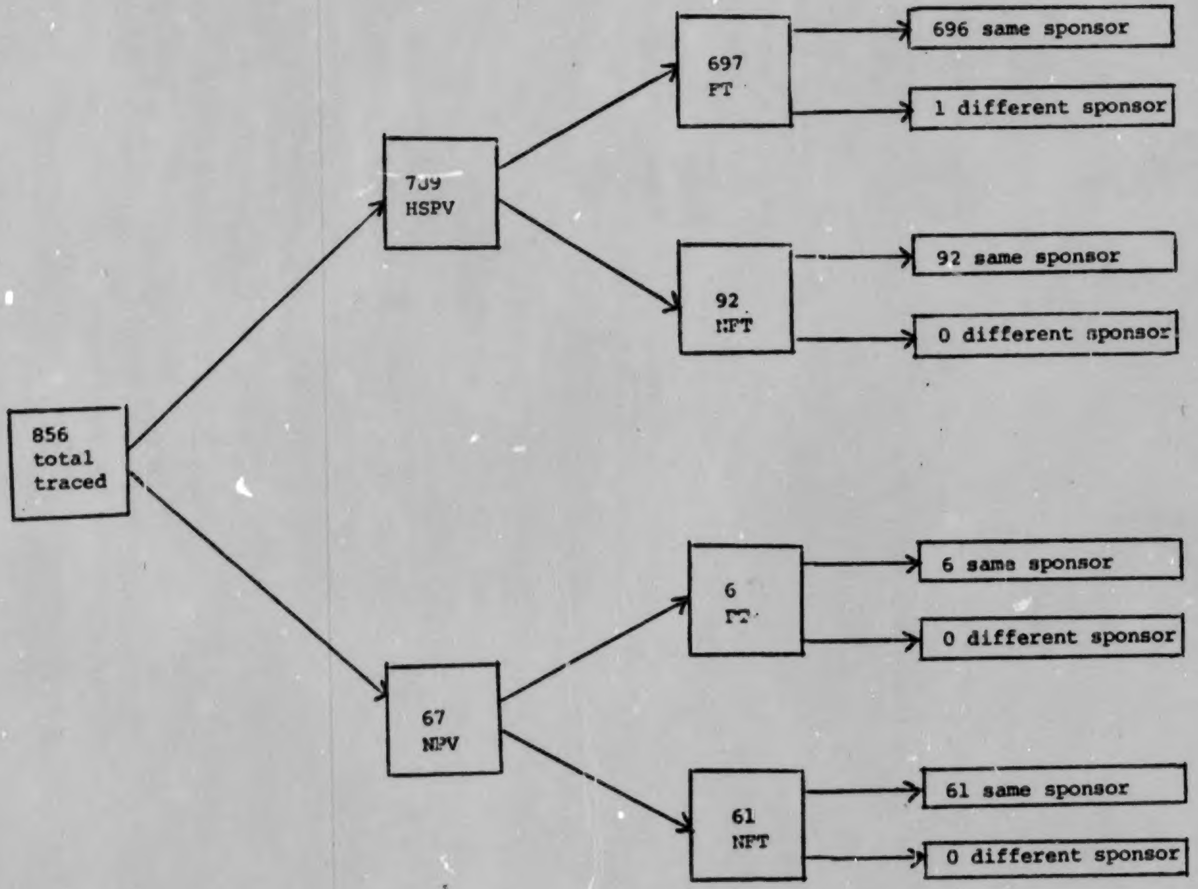

856 total traced
789 HSPV
67 NPV
697 FT
92 NFT
6 FT
61 NFT
696 same sponsor
1 different sponsor
92 same sponsor
0 different sponsor
6 same sponsor
0 different sponsor
61 same sponsor
0 different sponsor

Figure MIV-3   Treatment History Distribution of HSPV-FT Children in Cohort IV the basis for stable comparisons, even if all the traced children had taken all the tests in each battery. In fact, however, the testing pattern reduces the groups available for analysis even further, as appears in Figure MIV-4. Only 251 of the 416 traced Cohort III children received the full pretest battery in Head Start, and only 263 received the full post-test battery in spring of first grade. If we wish to examine progress from pretest to posttest by way of an intermediate test, we have only 139 or 152 cases to study, depending on our choice of a spring or fall intermediate test, respectively. A few of those, moreover, are NPV and NFT cases: too few for comparison, but enough to reduce our "samples" still further.

Given these severely restricted data resources, we clearly cannot draw definitive conclusions about the value of continuous sponsorship in compensatory education. We can--and do--for Cohort III only:

- describe the data in some detail;
- look for some of the effects that the theory of HSPV-FT might lead us to expect; and
- study the patterns of sponsor and site differences within the HSPV-FT treatment population, without reference to "non-treatment" groups.

2.2 Background Characteristics

Figure MIV-4 displays various background characteristics of the overlapping groups of children tested on the four testing occasions: Fall Head Start (FH), Fall and Spring Kindergarten (FK and SK, respectively), and Spring First Grade (S1). It shows also the same characteristics of three longitudinal subsets of children tested in Head Start and then on the three possible pairs of the remaining occasions, as illustrated in Figure MIV-5.

No clear demographic trends emerge across the four testing occasions: HSPV-FT seems to have tested comparable groups of children each time.

Subset of Traced Children	N	% Female	% Black	% White	Mean Mother's Education	Mean Age (as of 10/1/71)	Mean Household Size	Mean Income
Tested Fall Head Start (FH)	251	45%	34%	52%	10.3	1986	5.54	$3623
Tested Fall Kindergarten (FK)	372	44%	28%	60%	10.3	1990	5.46	$3794
Tested Spring Kindergarten (SK)	331	44%	30%	57%	10.2	1990	5.61	$3840
Tested Spring First Grade (SI)	263	47%	34%	56%	10.4	1989	5.53	$3761
Tested FH, FK, and SK	188	46%	36%	51%	10.2	1985	5.66	$3619
Tested FH, FK, and SI	152	50%	40%	50%	10.4	1985	5.67	$3598
Tested FH, SK, and SI	139	51%	42%	49%	10.2	1988	5.74	$3563
All Traced Children	416	44%	28%	58%	10.3	1992	5.47	$3809

Figure MIV-4 Background Variables for Traced Children and for Various Tested Subsets of Traced Children in Cohort III.

	"Short" (2 years traced)	"Long" (3 years traced)
"Early" (intermediate test in the fall)	FHFKSK	FHFKS1
"Late" (intermediate test in the spring)	——	FHSKS1

Figure MIV-5 Structure of Three-Point Longitudinal System

The three longitudinal traced samples do show some slight but interpretable trends in the characteristics of children who received the full sequence of three test batteries:

- The longitudinal samples are slightly poorer and younger, have more black and female members, and come from larger families than the testing-occasion samples of which they are the intersections.
- The "later" and "longer" samples are likewise slightly poorer and younger, have more black and female members, and come from larger families than the "earlier" and "shorter" samples.

Figure MIV-5 illustrates these distinctions.

There seems to be a slight trend toward greater attendance rates among black females than among the other categories of children over the time span of this study. It is difficult to attribute this trend to the cultural status of black females (which might serve to keep them in the same school situation, thereby allowing them to be tested and traced), rather than to the peculiarities of the testing and tracing procedures. This is a point to be examined further in future studies.

Figure MIV-6 disaggregates the background characteristics of the site level for each testing occasion. Again, no interesting trends emerge across testing points within sites, but we find the usual pattern of high

Figure MIV-6

Site Means of Background Variables at Four Testing Points, Cohort II Traced Children

Sponsor	Site	Testing Point	N	% female	% black	% white	Mother's Education	Age (days as of 10/1/71)	Household Size	Income
Far West	Duluth	FH	30	37	13	80	--	2,032	5.5	4,571
		FK	79	38	5	92	--	2,019	5.3	4,938
		SK	73	41	7	89	--	2,019	5.4	5,020
		SI	55	40	6	94	--	2,027	5.1	4,692
	Tacoma	FH	25	36	24	68	11.2	2,006	5.3	3,417
		FK	52	38	[illegible]	67	10.9	2,005	5.3	3,728
		SK	26	38	27	69	10.4	2,017	5.7	3,644
		SI	34	38	29	62	11.0	2,010	5.3	3,818
Arizona	Lincoln	FH	37	38	14	78	10.9	1,981	5.3	3,985
		FK	68	34	12	79	10.7	1,979	5.6	3,924
		SK	61	36	15	75	10.5	1,989	5.8	3,956
		SI	49	39	20	69	10.8	1,975	5.8	3,926
Kansas	Portage-ville	FH	55	54	39	61	9.7	1,989	5.7	3,493
		FK	55	54	39	61	9.7	1,989	5.8	3,511
		SK	54	56	40	60	9.7	1,989	5.8	3,483
		SI	45	58	43	57	9.7	1,987	5.6	3,464
High/ Scope	Fort Walton Beach	FH	41	51	83	17	10.1		6.6	3,440
		FK	46	54	83	17	10.0	1,932	6.5	3,416
		SK	45	53	80	20	10.0	1,931	6.5	3,386
		SI	32	56	91	9	9.9	1,924	6.5	3,330
	Greeley	FH	38	42	0	29	9.6	2,010	5.2	3,582
		FK	38	40	0	21	9.4	2,020	5.0	3,642
		SK	45	42	0	29	9.8	2,014	5.2	3,869
		SI	23	52	0	26	9.6	2,029	5.8	3,991
	Seattle	FH	25	52	60	36	11.6	1,974	4.6	3,442
		FK	34	58	62	28	11.2	1,967	4.2	3,169
		SK	27	44	74	22	11.7	1,951	4.5	3,246
		SI	25	52	72	24	11.5	1,957	4.4	3,038

variability from site to site. A few outliers deserve attention here at the outset:

- Duluth is substantially richer and has a higher proportion of whites than any other site.
- Greeley's 75% Chicano population is the only appreciable non-black, non-white ethnic group in the sample.
- Only Seattle and Fort Walton Beach have a black majority.
- Fort Walton Beach has a higher proportion of black and young children than any other site.

2.3 Measures

Data collection procedures included the administration of several child tests as well as teacher and parent information forms to a complete or partial sample in each location.

Smith (1973) and Weisberg (1973) provide descriptions of all outcome measures used in the HSPV study, including quantitative reliability estimates. We describe here only those measures which are relevant to HSPV/FT.

1. Caldwell Preschool Inventory (PSI)

The PSI was designed especially for use with Head Start populations. Its purpose is to evaluate achievement of disadvantaged children in terms of skills required for success in school. Used as a pre- and post test, it is an index of educational achievement due to Head Start participation. It has been found useful for this purpose, having high reliability and high correlations with other cognitive tests. It does not appear to be too easy or too hard (no "ceiling" or "floor" effects) for the HSPV population, either during Head Start or in Kindergarten.

2. NYU Books 3D and 4A

These measures of pre-school achievement tap a variety of specific dimensions. Book 3D assesses pre-math and pre-science achievement and linguistic concepts, and Book 4A assesses knowledge of alphabet, numeral, and shape names. We include these in the group of pretests mainly for the sake of continuity with Smith, whose preschool effects studies used them extensively. They were administered only to the Cohort III children, and only during Head Start.

3. Peabody Picture Vocabulary Test (PPVT)

The PPVT is a vocabulary test in which the child is required to match one of four pictures to a word spoken by an examiner. Internal reliability is reasonably high. The test is neither too easy nor too hard for young children; in fact, since items increase in difficulty and the child continues until he makes six errors out of eight consecutive items, there is practically no upper score limit for young children.

While the test may tap more general types and levels of skills, we use it only as an index of passive vocabulary.

4. Wide Range Achievement Test (WRAT)

The WRAT consists of spelling, reading, and arithmetic subtests. These subtests measure specific academic skills in each domain. The spelling subtest includes mark copying items as well as single-syllable spelling items; this subtest purports to measure motor development and general school readiness. The Reading subtest taps visual discrimination and general knowledge skills in matching and recognizing letters and reading words. The Arithmetic subtest asks the child to count concrete objects, to read numbers, to perform simple addition and subtraction, and to distinguish between "more" and "less". Both written and verbal abilities are measured.

Each of these subtests measures specific skills, has reasonable reliability and correlates moderately highly with other tests. In the earlier HSPV studies, however, ceiling and floor effects raised some analytic problems.

5. Metropolitan Achievement Test (MAT)

The MAT was administered to Follow Through children from kindergarten on and therefore does not figure in the HSPV studies. The Primary Form for first grade use includes a Word Analysis subtest, a two-part (Word Knowledge and Reading) Total Reading Subtest, and a Mathematics subtest. The Word Analysis subtest requires the child to select a spoken word from a written list of similar words. The test requires auditory discrimination and phonic and linguistic skills.

The two Reading subtests require the child to apply a diverse array of skills as he identifies words matching each of a series of pictures, marks sentences which match pictures, and reads short paragraphs and answers questions about them. Sight vocabulary, phonic skills, elimination procedures, contextual clues, and comprehension are all used to arrive at correct answers.

The Mathematics subtest taps concept manipulation and computational skills. There is, however, only one score for the entire subtest.

In general, the subtests tap a wide variety of skills and measure achievement in several domains. The MAT test is a reliable instrument. It shows no ceiling or floor effects in the HSPV-FT study.

2.4 The Logic of the Two HSPV-FT Studies

Against the backdrop of the data situation described above, we now present the results of two studies which exploit the positive capabilities of the HSPV-FT Cohort III data base.

Study I: Kindergarten and First Grade Academic Correlates of HSPV and Other Preschool Experiences

This first study comes as close as the data permit to comparing effects of experience in HSPV with those of other kinds of preschool experience. It asks whether children who have been through HSPV perform better as a group in kindergarten and first grade than others who had unsponsored Head Start or no formal preschool experience.

Study II: Patterns of Site Means and Ranks on Academic and Affective Measures among HSPV-FT Children.

Study II is a descriptive study of the test scores of the children who received the pretest battery in the HSPV treatment population and then the posttest battery as FT children at the end of first grade. It examines the mean score differences from site to site and the shifting patterns of site ranks over time for the diverse academic and affective measures.

3.0 STUDY I: KINDERGARTEN AND FIRST GRADE ACADEMIC CORRELATES OF HSPV AND OTHER PRESCHOOL EXPERIENCES

Does continuity of curricular approach from Head Start through kindergarten and first grade improve the school performance of participant children? Given the data at hand, we can attempt only an indirect approach to this question. The best we can do is to carry out a posttest-only comparison of certain nonrandomized comparison groups with our traced set of children, each of whom took part in Head Start and Follow Through under one of four Sponsors. We have no way of taking quantitative account of the many confounding ways in which the HSPV-FT group within each Sponsor would have differed from the comparison groups, even if HSPV-FT had never happened.

The most reasonable summary of the analyses described in the remainder of this chapter is that continuous HSPV-FT experience is not associated with academic performance at the end of kindergarten or at the end of first grade. These results are far from conclusive, however, and the character of the data forces us to withhold judgment on the efficacy of the continuous experience when examined with these data.

3.1 The Sample

The Follow Through children who had HSPV preschool experience were traced into the larger analytic sample of all FT and NFT subjects. Within the sites where HSPV children showed up, we formed four post hoc "treatment" groups:

- HSPV-FT
- Home-FT
- HS-NFT (some unspecified Head Start experience)
- Home-NFT.

Figure MIV-7 indicates the distribution of subjects among treatment groups within the seven sites under four Sponsors; it displays also the group means of annual family income, mothers' education, and reading and math scores, standardized with a mean of 50 and a standard deviation of 10 relative to the entire NFT sample at kindergarten entry (FK), at the end of

Figure MIV-7 Background Characteristics and Mean Test Scores of Children by Site and Treatment Group

Sponsor	Site	Treatment Group	N	Income	Mother's Education	Fall Kindergarten Pretests WRAT		Spring Kindergarten Tests WRAT		Spring Grade 1 Tests MAT	
						Reading	Math	Reading	Math	Reading	Math
Far West	Duluth	HSPV-FT	41	4750	9.9	51.0	53.0	51.1	53.1	48.5	51.9
		HOME-FT	42	7091	10.4	53.4	52.5	53.8	54.2	51.3	53.3
		HS-NFT	17	5724	8.3	49.3	53.5	48.2	49.2	49.5	47.5
		HOME-NFT	52	8016	11.8	53.7	56.1	54.4	57.8	55.3	54.0
	Tacoma	HSPV-FT	11	4546	9.9	52.2	52.8	51.4	52.6	43.0	43.9
		HOME-FT	10	6668	13.8	53.8	52.5	53.1	51.8	51.4	48.8
		HS-NFT	2	2850	8.0	49.0	54.9	45.6	53.3	47.4	50.7
		HOME-NFT	12	7925	10.2	57.5	57.5	56.9	58.3	54.9	59.0
Arizona	Lincoln	HSPV-FT	28	4600	9.8	49.1	50.3	49.0	49.1	46.9	48.1
		HOME-FT	110	7200	11.9	55.0	54.2	56.0	53.3	53.2	52.4
		HS-NFT	9	6560	9.7	48.8	46.0	43.9	49.3	44.1	48.6
		HOME-NFT	83	8040	13.1	56.7	56.9	54.3	55.3	58.1	57.7
Kansas	Portage-ville	HSPV-FT	34	4558	8.2	52.8	52.7	55.1	57.0	50.5	55.3
		HOME-FT	20	7376	10.4	48.1	51.3	55.4	60.6	56.3	59.3
		HS-NFT	23	3109	8.0	42.0	44.0	45.6	47.7	51.1	50.1
		HOME-NFT	17	5840	8.0	46.0	49.2	51.3	55.1	60.3	57.2
High/ Scope	Fort Walton	HSPV-FT	26	3933	8.4	42.7	43.5	40.0	43.9	42.5	37.3
		HOME-FT	16	4214	9.8	42.3	42.7	40.6	43.0	48.2	37.8
		HS-NFT	4	4800	8.0	44.4	42.4	48.7	47.5	50.6	48.5
		HOME-NFT	48	5525	8.4	45.4	44.7	43.6	42.3	48.2	49.9
	Greeley	HSPV-FT	17	4748	8.0	49.9	54.0	40.8	42.6	45.1	48.3
		HOME-FT	9	7568	9.0	55.2	54.5	53.4	59.1	52.2	54.0
		HS-NFT	12	5034	8.1	51.6	55.8	49.4	54.4	45.9	50.6
		HOME-NFT	6	7234	8.0	50.8	54.6	46.4	52.0	47.9	49.0
	Seattle	HSPV-FT	12	2167	12.0	47.2	50.0	58.7	50.9	42.1	40.1
		HOME-FT	23	4767	10.1	54.6	50.2	54.1	53.1	46.8	45.1
		HS-NFT	2	7000	12.0	66.0	62.6	61.2	57.2	58.8	61.8
		HOME-NFT	32	7346	11.8	53.3	50,8	50.8	48.1	50.1	49.1

kindergarten (SK), and at the end of first grade (S1). A few patterns deserve mention here:

- Group mean income varies rather more within sites (among treatment groups) than among sites; some of this variability stems, no doubt, from the small sizes of some groups.
- In five of the seven sites, the HSPV-FT group has the lowest annual family income of the four treatment groups. In five sites, moreover, the Home-NFT group has the highest family income.
- Each site seems to have its own fairly characteristic mean levels of mothers' education, with Greeley, Fort Walton, and Portageville reporting generally less-schooled mothers than the other sites. The within-site ranks of the four treatment groups on this background variable follow no obvious pattern.
- On all the test scores, moreover, we note substantial variability among sites and among treatment groups within sites.

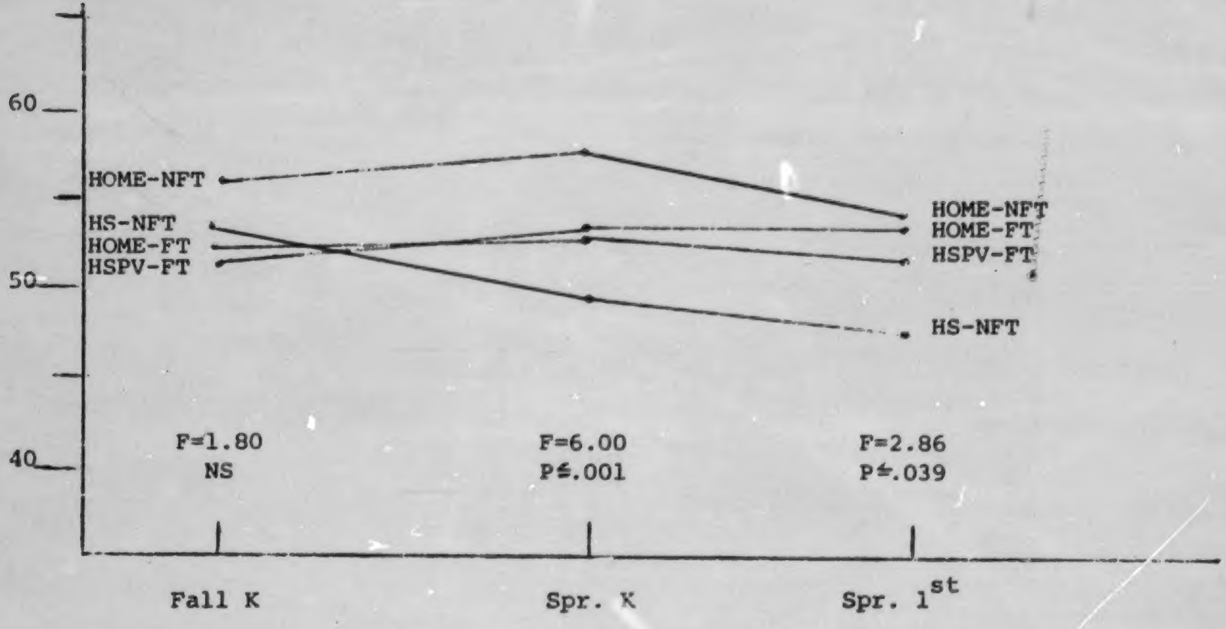
3.2 Longitudinal Performance Patterns for Treatment Groups Within the Seven Sites

Figure MIV-8 plots the reading and math standard scores tabulated in Figure MIV-7 across the three testing points for each treatment group in each site. At each time point, we have tested the variability of the treatment group means with a one-way analysis of variance; the resulting F-ratio and its associated probability level appear at each breaking point on each graph. Two groups with only two members each were omitted. Note that the contents of the tests in each domain vary somewhat across the three testing occasions, as of course they must in order to be appropriate to the changing developmental level of the children.

Figure MIV-9 summarizes the general shapes of these broken curves for the Follow Through groups.

These displays provide little evidence that HSPV experience stimulated superior progress in kindergarten and first grade. Such evidence as we have suggests that any academic gains appeared early and tended to dissipate. The performance of HSPV children in kindergarten and first seems indistinguishable, on the basis of the evidence at hand, from that of children who stayed home before kindergarten. While some kindergarten gains relative to NFT emerge, no HSPV-FT or Home-FT group gained in first grade. The bulk of the changes were either non-significant or negative. The data give us no firm basis for discerning any difference between the performance of HSPV and "home" as preschool treatments.

Math



Reading

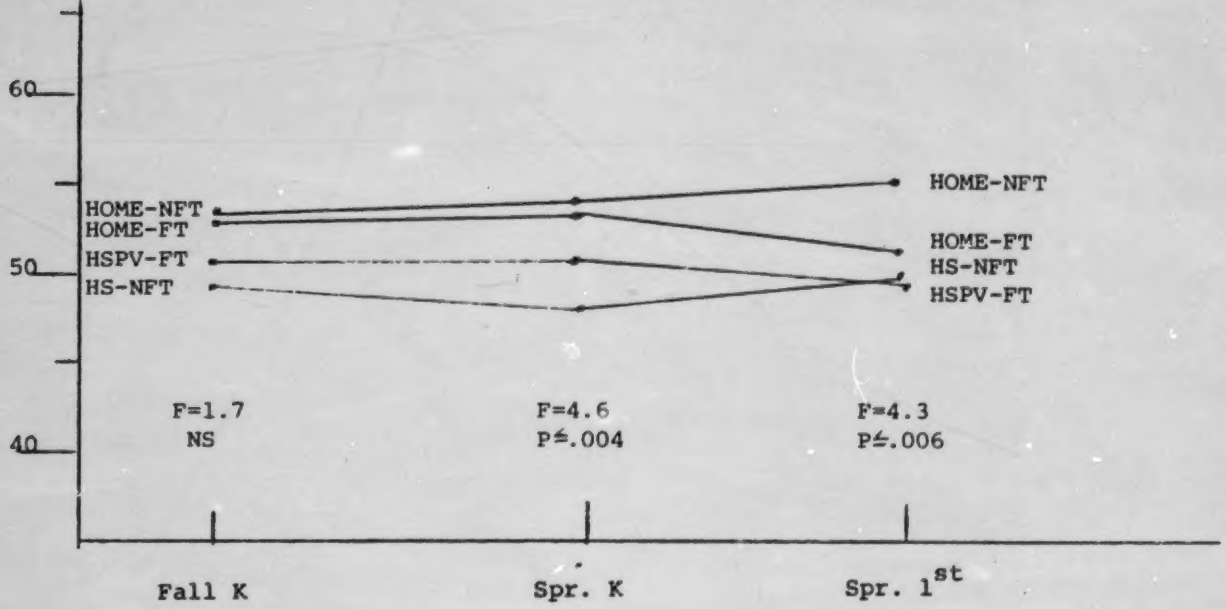
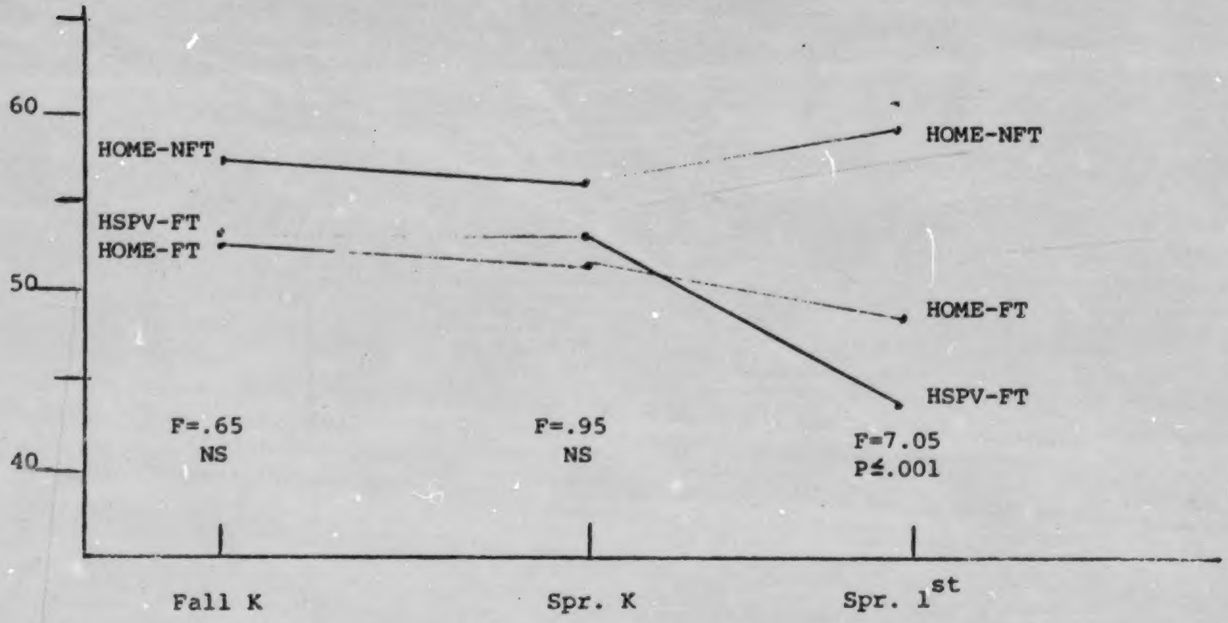


Figure MIV-8a

Treatment Group Means of Math and Reading Scores for Duluth

Math



Reading

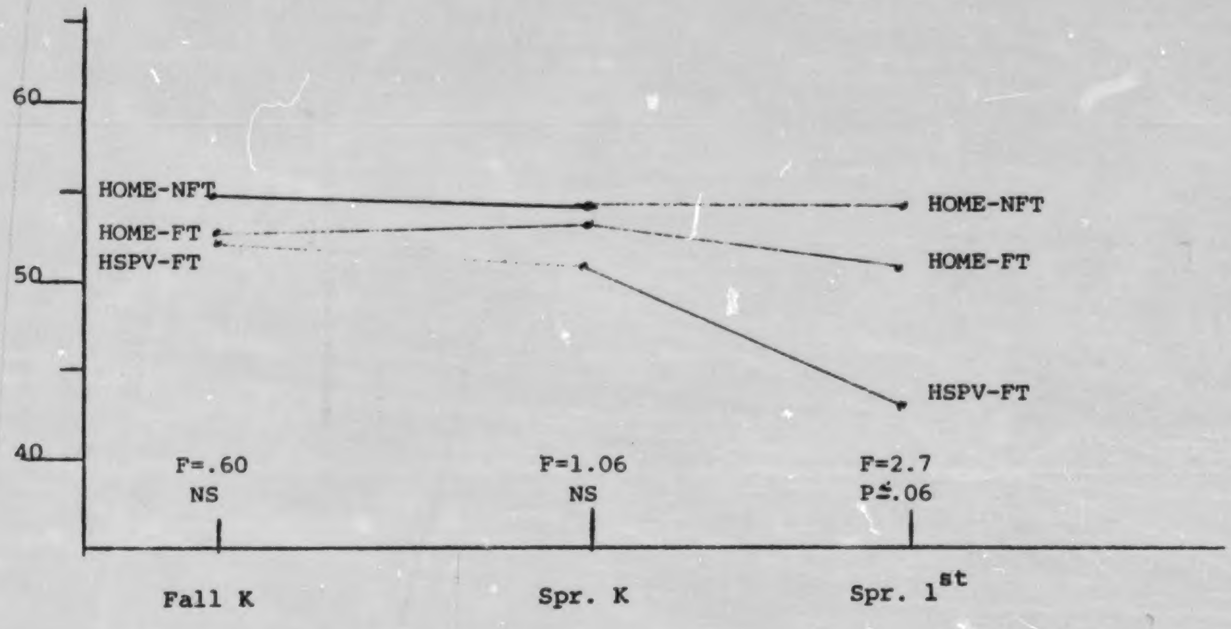
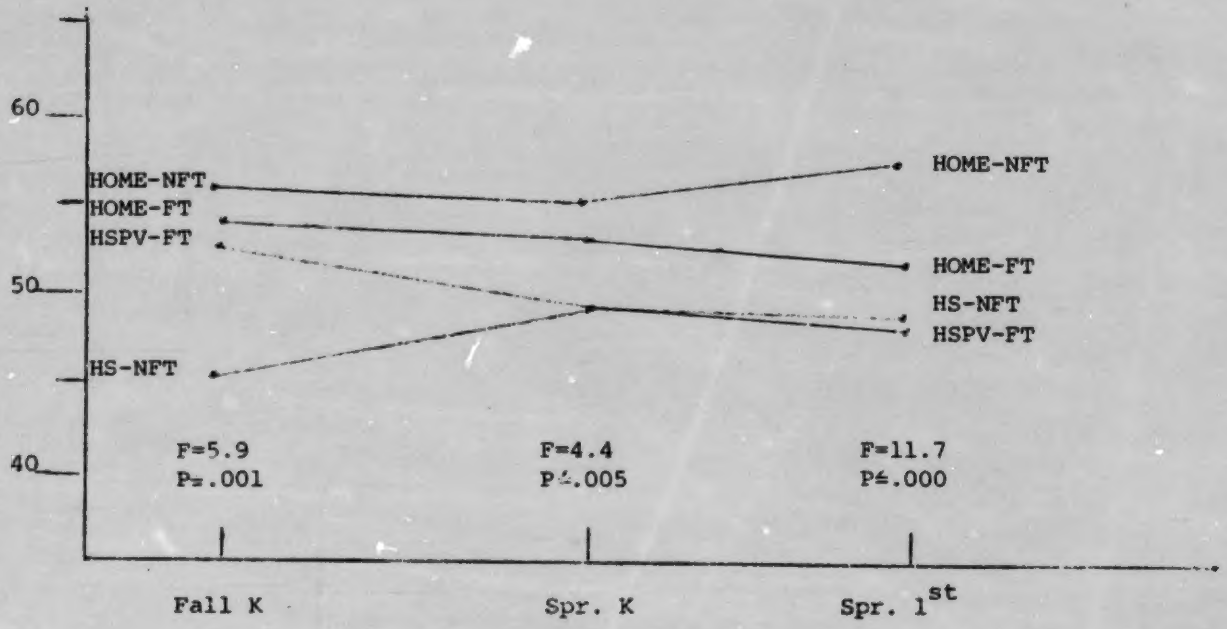


Figure MIV-8b

Treatment Group Means of Math and Reading Scores for Tacoma

00237

Math



Reading

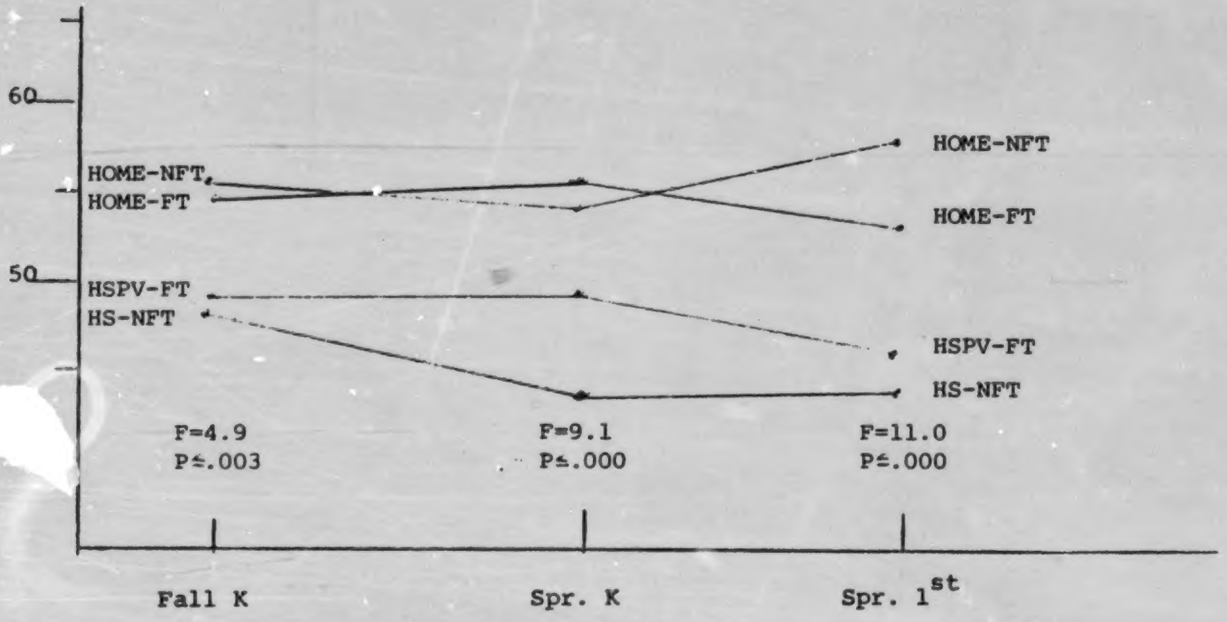
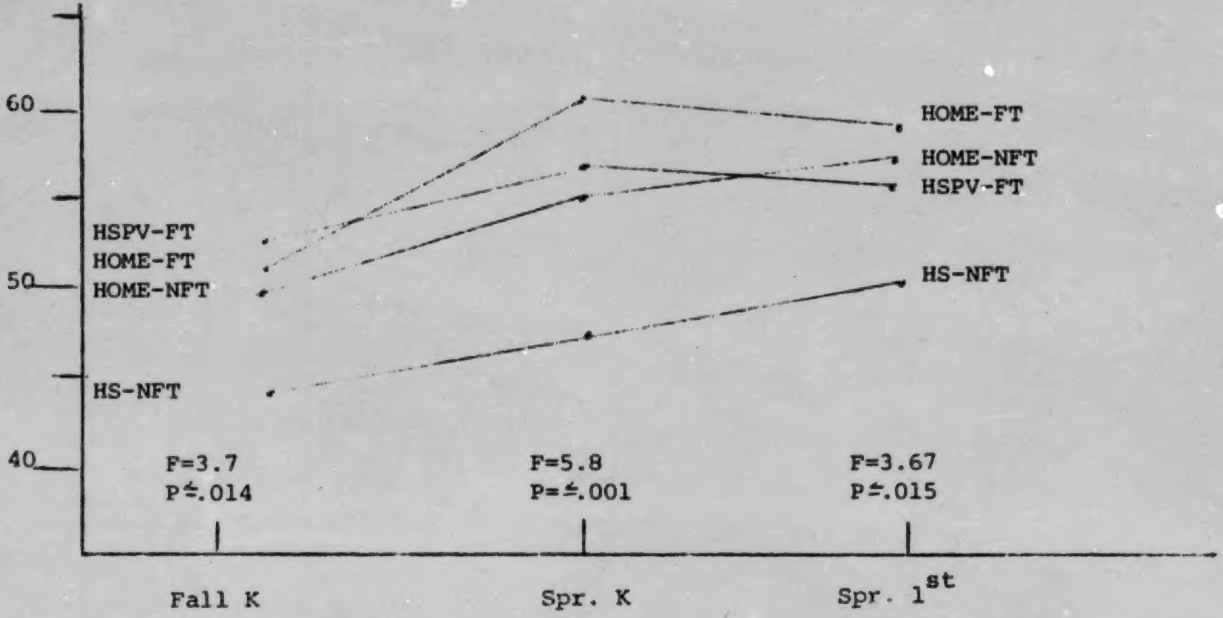


Figure MIV-8c

Treatment Group Means of Math and Reading Scores for Lincoln

Math



Reading

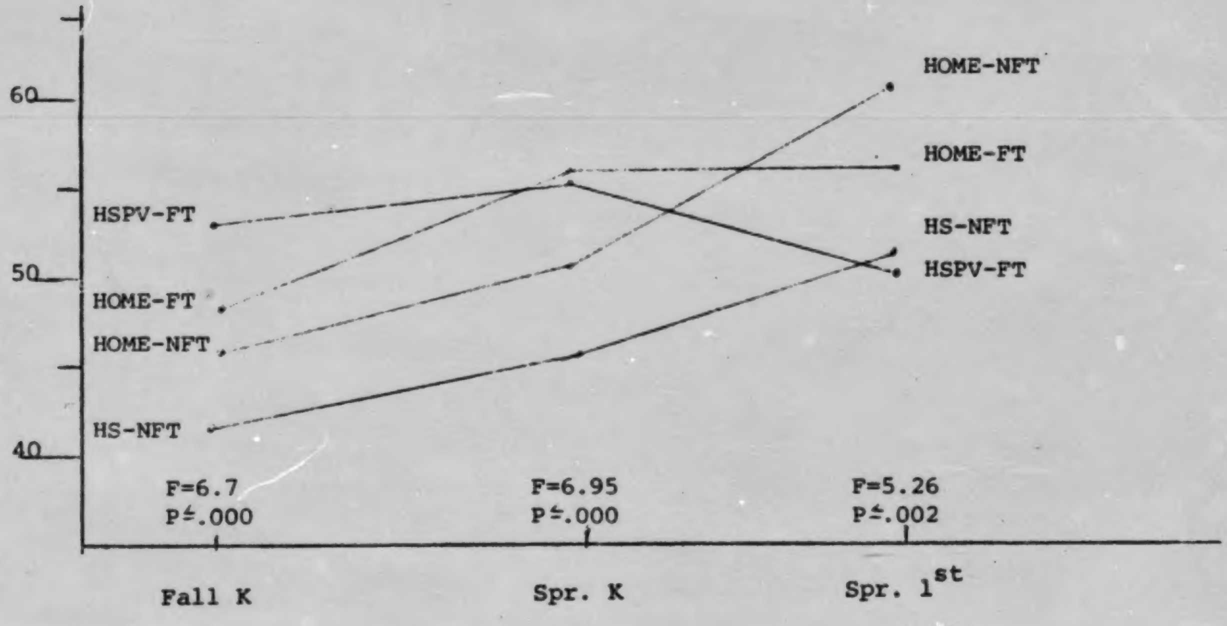
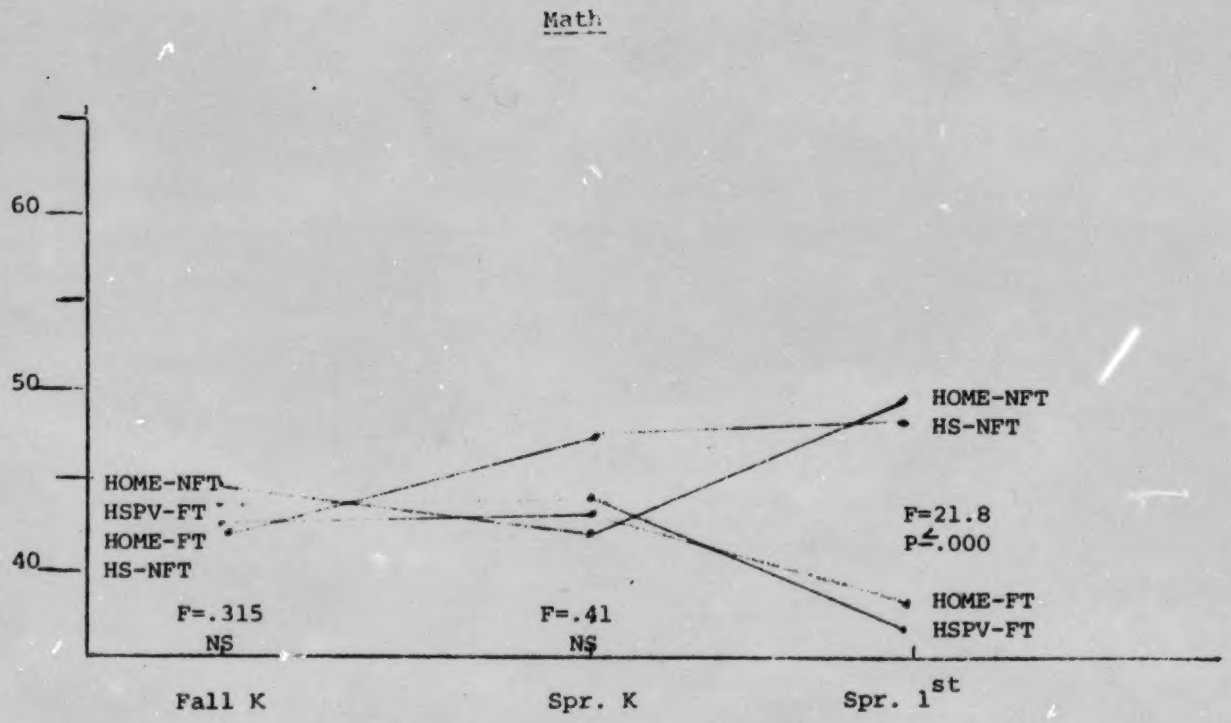


Figure MIV-8d

Treatment Group Means of Math and Reading Scores for Portageville


Math

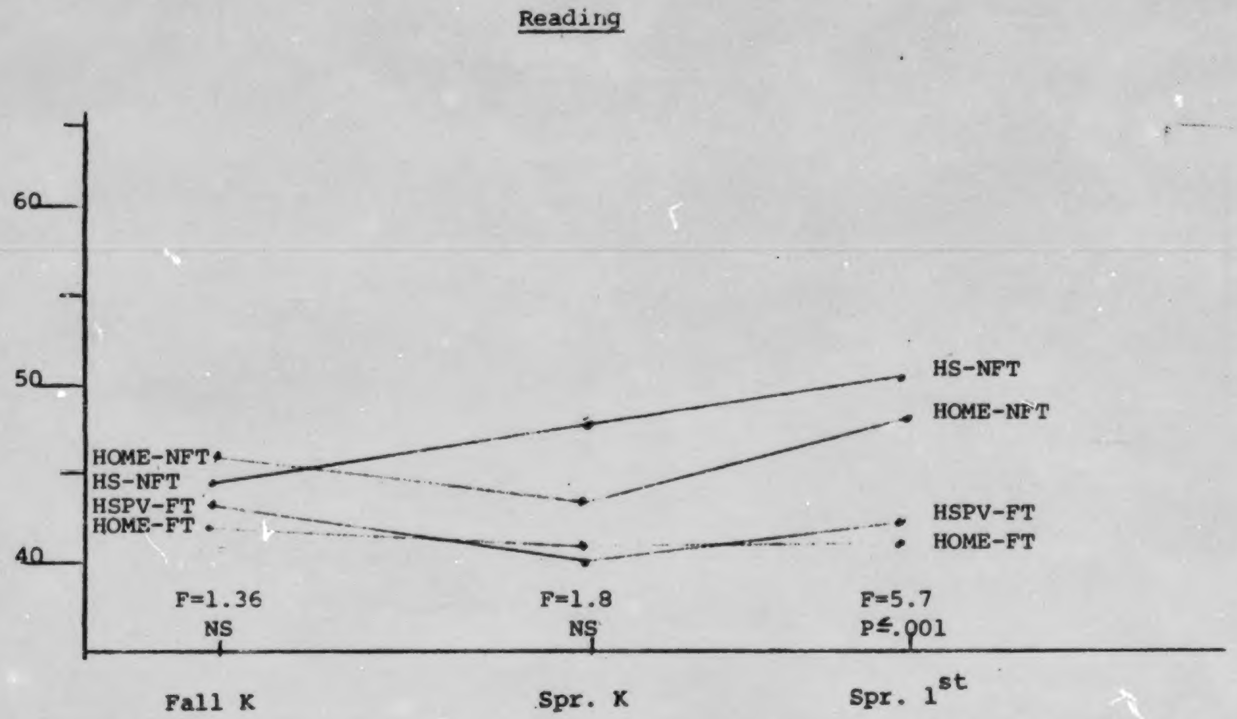

Reading

Figure MIV-8e

Treatment Group Means of Math and Reading Scores for Fort Walton Beach

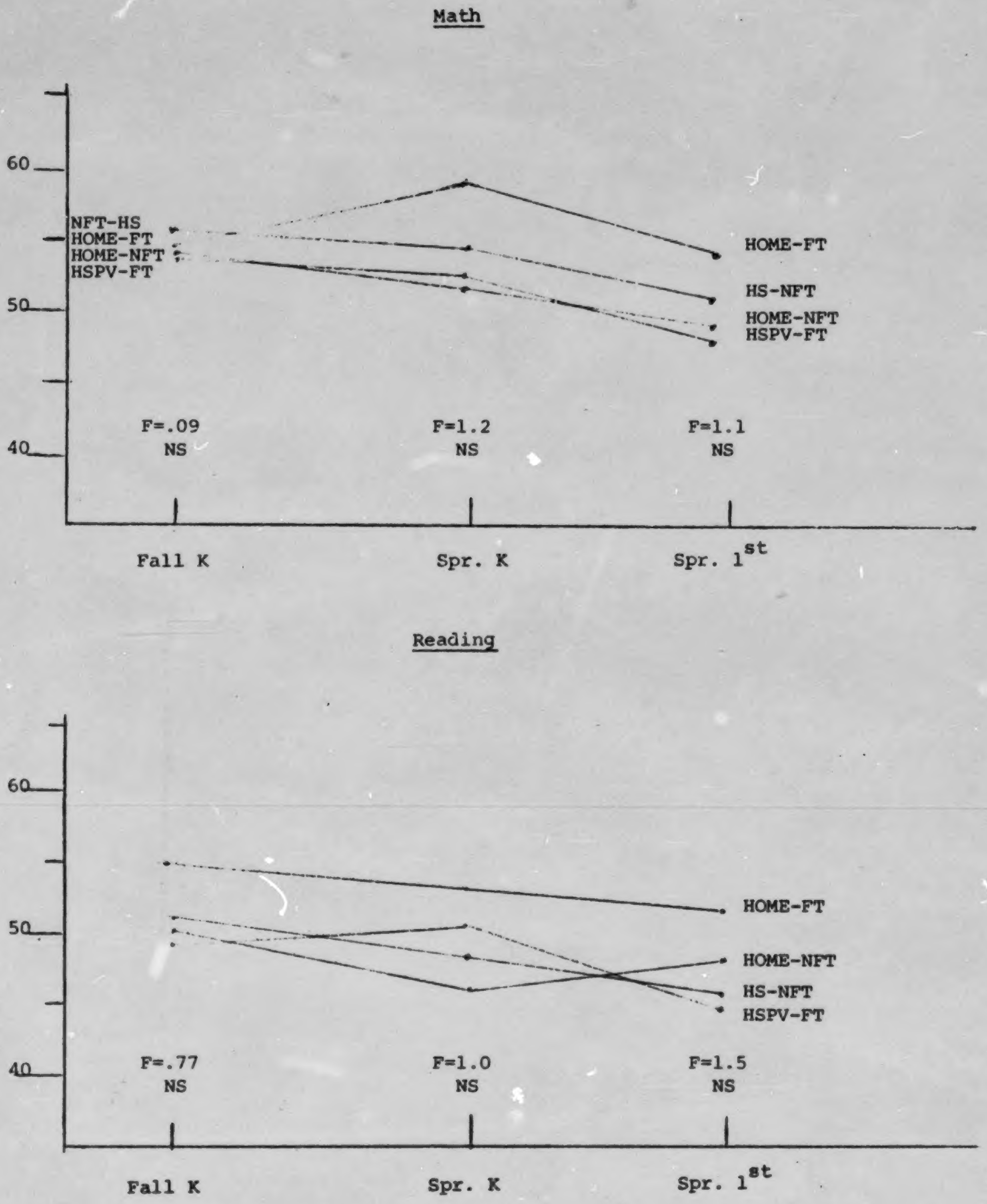
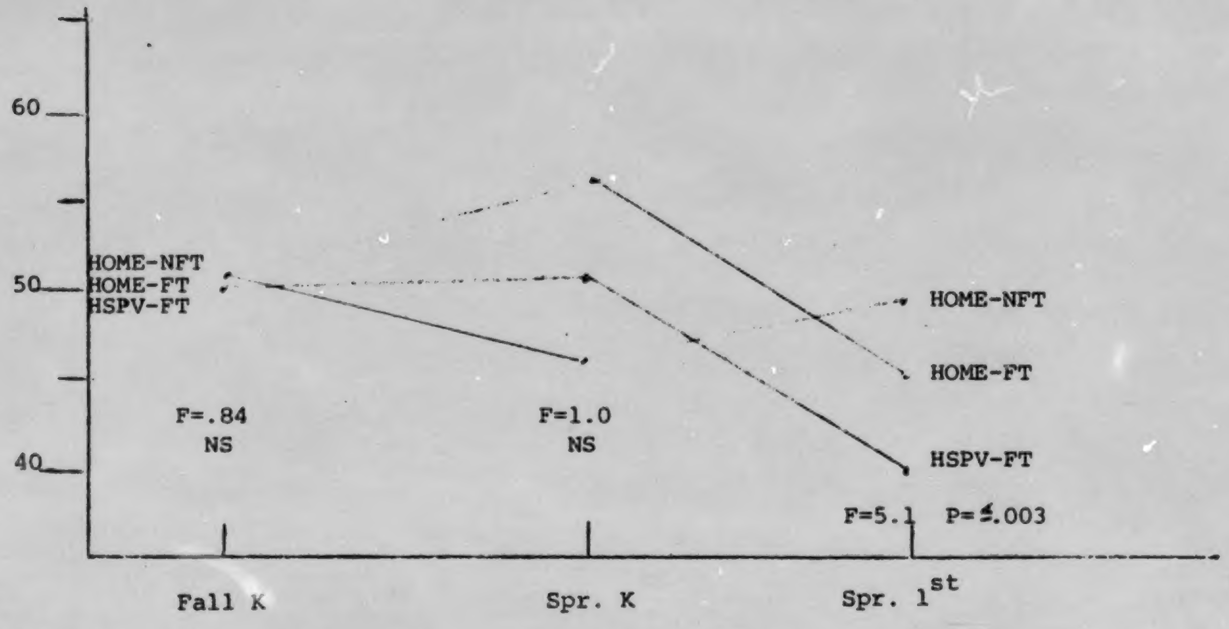
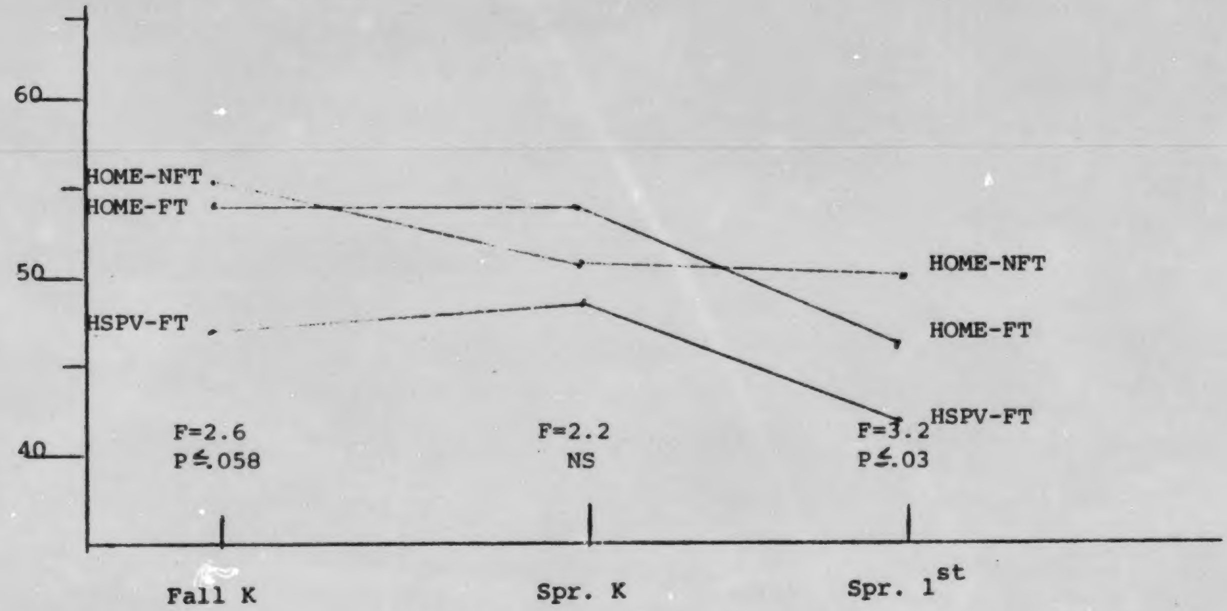


Figure MIV-8f

Treatment Group Means of Math and Reading Scores for Greeley



Math

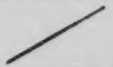
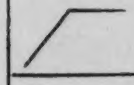
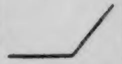
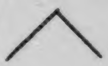
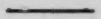
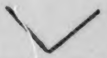
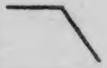
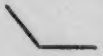
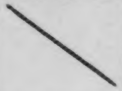


Reading

Figure MIV-8g

Treatment Group Means of Math and Reading Scores for Seattle

Figure MIV-9 Longitudinal Performance Patterns for FT Groups Within Sites

Pattern		HSPV		HOME	
		Reading	Math	Reading	Math
	K gain 1st gain				
	K gain 1st even		Portageville	Portageville	Portageville
	K even 1st gain				
	K gain 1st loss	Greeley Portageville			
	K even 1st even	Duluth Fort Walton	Duluth Lincoln	Duluth Fort Walton	Duluth Lincoln
	K loss 1st gain				
	K even 1st loss	Tacoma Lincoln	Tacoma Fort Walton Seattle	Tacoma Lincoln Seattle	Tacoma Fort Walton Seattle
	K loss 1st even				
	K loss 1st loss		Greeley	Greeley	

4.0 STUDY II: PATTERNS OF SITE MEANS AND RANKS ON ACADEMIC AND AFFECTIVE MEASURES AMONG HSPV-FT CHILDREN

Even though we cannot ask about the absolute effects of continuous sponsorship through Head Start and Follow Through, we can still ask the data how the various Sponsors, under the distinctive conditions of their diverse sites, performed relative to each other. In the context of the ongoing Follow Through analytic studies, do the HSPV-FT data case any useful light on the Sponsors' relative abilities to stimulate specific kinds of progress in specific sorts of children under specific circumstances?

4.1 The Sample

Figure MIV-10 displays for each site the background characteristics of the set of HSPV-FT (treatment only) children with both pretest and posttest information. These 137 children, distributed as shown in Figure MIV-4, display the same kinds of intersite variability that Section 2.2 describes. By requiring both pretest and posttest, of course, we reduce the Study II sample to a much smaller group than that which we examined in Study I.

4.2 Outcomes

Figure MIV-11 displays site means and ranks of seven pretests administered at the beginning of Head Start, 14 intermediate tests administered during kindergarten, and the various MAT posttests administered at the end of first grade. Figure MIV-12 summarizes the site ranks for the four testing occasions, separating out the Spring kindergarten affective measures as a separate category.

The pretest ranks parallel each other quite closely, and we can therefore speak meaningfully of an overall dimension of "entry level" which varies quite clearly from site to site: Greeley starts out above all the other sites, with Seattle, Tacoma, and Duluth close together and not too far behind. Lincoln and Portageville occupy a substantially lower level, and Fort Walton Beach seems almost to belong on a different distribution. Let us note in passing that these entry level data illustrate the pervasive superiority of site-level over Sponsor-level analyses in Follow Through. High/Scope as a Sponsor faces radically different challenges in Greeley and Seattle from those in Fort Walton Beach: to aggregate the three sites is necessarily to create a fuzzy picture of each such situation.

Figure MIV-10 Site Means of Background Variables for HSPV-PT Children Tested in Fall of Head Start anć in Spring of Grade 1

Sponsor	Site	N	% female	% black	% white	mothers' education	age (in days as of 10/1/71)	household size	annual income
Far West	Duluth (204)	17	35	12	88	--	2043	5.9	4333
	Tacoma (213)	15	40	33	60	11.7	2018	5.4	3814
Arizona	Lincoln (316)	13	46	15	85	11.4	2006	5.6	4603
Kansas	Portage-ville (804)	39	56	45	55	9.6	1996	5.5	3569
High/ Scope	Fort Walton Beach (902)	26	50	92	8	10.0	1924	6.8	3436
	Greeley (906)	13	69	0	38	9.8	2008	6.2	3931
	Seattle (910)	14	50	57	36	11.8	1958	4.5	3000
All Children		137	50	43	50	10.4	1989	5.8	3672

Figure MIV-11a Site Means and Ranks of Means of Pretest Variables for Children Traced from Pretest (Head Start) to Postest (Spring, Grade 1)

Sponsor	Site	N	Book 3D		Book 4A				Pre-School Inventory							
					Alphabet		Numbers		Social Responsiveness		Associate Vocabulary		Numerical		Sensors	
			Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
Far West	Duluth	17	12.88	6	2.65	4	3.18	5	11.12	3	6.41	5	7.35	5	10.94	4
	Tacoma	15	12.00	5	2.93	6	3.20	6	12.53	5	6.40	4	7.00	4	12.67	5
Arizona	Lincoln	13	10.46	3	2.31	3	1.77	2	11.15	4	4.62	2	5.62	3	9.69	2
Kansas	Portage-ville	39	10.38	2	2.28	2	2.49	3	10.36	2	4.97	3	5.26	2	10.13	3
High/ Scope	Fort Walton	26	8.23	1	1.23	1	1.54	1	7.46	1	2.19	1	3.81	1	8.35	1
	Greeley	13	13.46	7	2.85	5	4.00	7	13.69	7	7.77	7	8.85	7	15.23	7
	Seattle	14	11.64	4	3.50	7	2.93	4	12.79	6	7.43	6	7.93	6	13.36	6
All Children		137	10.89		2.38		2.59		10.78		5.26		6.08		10.94	

Figure MIV-11b Site Means and Ranks of Means of Fall Kindergarten Test Scores for Children Traced from Pretest to Postest

Sponsor	Site	N	Pre-School Inventory						PPVT		WRAT					
			Responsiveness & Vocabulary		Numerical Concepts		Sensory Concepts				Math		Reading		Spelling	
			Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
Far West	Duluth	17	8.29	5	5.59	4	5.12	5	49.65	4	16.76	6	14.94	7	5.71	5
	Tacoma	15	7.27	2	6.13	7	5.20	6	49.07	3	15.80	4	14.60	5	3.53	2
Arizona	Lincoln	13	8.00	4	4.54	3	4.62	2	50.77	6	15.00	2	12.85	2	4.00	3
Kansas	Portage-ville	39	7.36	3	4.41	2	4.87	3	42.16	2	15.37	3	14.30	6	7.82	7
High/ Scope	Fort Walton Beach	26	6.23	1	4.23	1	4.38	1	33.23	1	10.52	1	9.00	1	1.96	1
	Greeley	13	9.31	7	5.62	5	5.08	4	52.00	7	16.58	5	14.33	4	4.67	4
	Seattle	14	9.29	6	6.00	6	5.57	7	50.46	5	18.71	7	14.29	3	6.64	6
All Children		137	7.69		5.00		4.91		44.71		15.11		13.45		5.20	

Figure MIV-11c Site means and ranks of means of spring kindergarten test scores for children traced from pretest to posttest.

Sponsor	Site	Locus of Control				Gumpgookies		PPUT		WRAT					
		Positive		Negative						Math		Reading		Spelling	
Far West	Duluth	5.88 n=17	4	4.53 n=17	3	49.3 n=17	5	55.6 n=17	5	23.8 n=17	5	21.7 n=17	5	9.94 n=17	6
	Tacoma	4.67 n=6	1	3.83 n=6	2	44.7 n=6	2	55.8 n=6	6	18.8 n=6	2	17.3 n=6	2	6.67 n=6	2
Arizona	Lincoln	6.08 n=21	5	3.77 n=21	1	38.7 n=21	1	54.9 n=21	3	21.3 n=21	3	20.1 n=21	3	9.85 n=21	5
Kansas	Portage-ville	5.79 n=43	3	4.58 n=43	4	49.7 n=43	6	50.3 n=43	3	24.3 n=43	7	24.0 n=43	7	12.29 n=43	7
High/ Scope	Fort Walton Beach	5.19 n=27	2	4.77 n=27	6	45.4 n=26	4	48.2 n=25	1	18.5 n=27	1	15.0 n=27	1	6.65 n=27	1
	Greeley	6.31 n=16	6	5.31 n=16	7	45.2 n=16	3	57.5 n=16	7	22.4 n=16	4	22.2 n=16	6	9.53 n=16	4
	Seattle	6.90 n=11	7	4.50 n=11	5	50.9 n=11	7	55.5 n=11	4	23.9 n=11	6	20.8 n=11	4	8.00 n=11	3
All Children		5.80 n=141		4.56 n=141		47.0 n=140		52.6 n=139		22.2 n=141		20.6 n=141		9.60 n=141	

Figure MIV-11d Site Means and Ranks of Means of Spring First Grade Test Scores for Children Traced from Pretest to Posttest

Sponsor	Site	N	MAT								Total	
			Word Knowledge		Reading		Math		"Total Reading"			
			Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
Far West	Duluth	17	44.8	5	19.5	6	34.8	6	40.4	6	99.1	6
	Tacoma	15	33.7	2	12.2	2	20.3	2	27.5	2	66.2	2
Arizona	Lincoln	13	44.2	4	17.2	5	25.7	4	37.1	4	87.1	4
Kansas	Portage-ville	39	46.1	6	19.8	7	36.6	7	41.1	7	102.5	7
High/Scope	Fort Walton Beach	26	25.4	1	12.15	1	13.2	1	24.7	1	50.8	1
	Greeley	13	53.0	7	14.7	4	28.8	5	38.8	5	96.5	5
	Seattle	14	36.6	3	13.1	3	21.9	3	29.4	3	71.6	3
All Children		137	40.2		16.1		26.9		34.6		83.2	

Figure MIV-12 Site Mean Ranks for Five Categories of Tests

Sponsor	Site	Pretests (FH)	Fall Kinder-garten (FK)	Spring Kinder-garten (SK) Academic	Spring Kinder-garten (SK) Affective	Spring 1st Grade (S1) Posttests
Far West	Duluth	4	5.5	5.5	3.5	6
	Tacoma	5	4	2	1	2
Arizona	Lincoln	3	2	3	2	4
Kansas	Portage-ville	2	3	7	5	7
High/ Scope	Fort Walton Beach	1	1	1	3.5	1
	Greeley	7	5.5	5.5	6	5
	Seattle	6	7	4	7	3

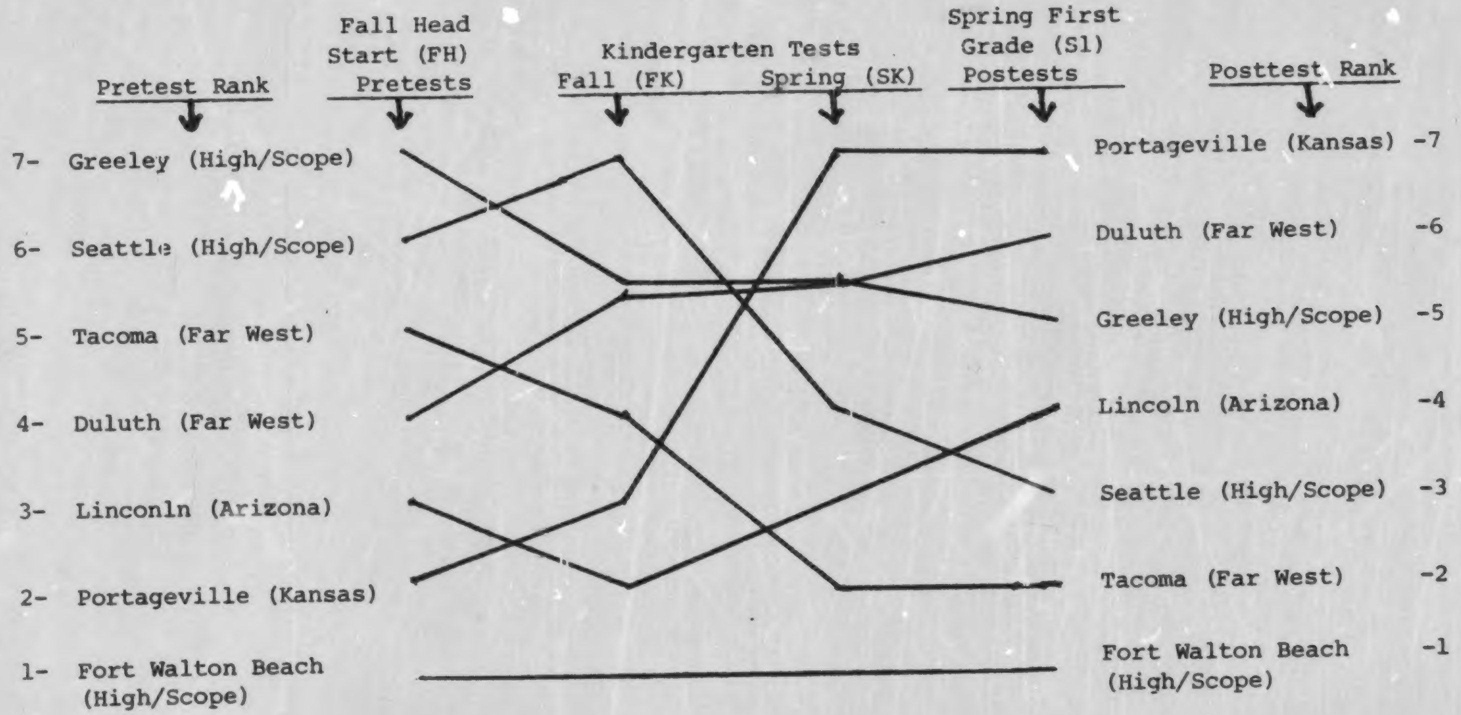


Figure MIV-13 Longitudinal Patterns of Mean Site Ranks of Academic Tests Scores over Four Testing Occasions.

By the time of intermediate testing a year later, as Figure MIV-12 and Figure MIV-13 show, Seattle, Greeley, and Duluth cluster together at the top of the rankings, with Tacoma, Portageville, and Lincoln in a lower band, and Fort Walton Beach completely out of the picture, particularly in its PPVT and WRAT scores.

The greatest overall changes in site rankings take place during kindergarten when:

- Portageville leaps from near the bottom to the top rank, consistently with our finding elsewhere that Kansas performs particularly strongly in kindergarten;
- Seattle falls from the top of the ranking to the middle; and,
- Tacoma continues its slide into the very low position that Portageville occupied at entry.

During first grade, no site changes its overall rank by more than a single position, and the posttest positions are approximately the same as those at the end of kindergarten. As at pretest, each site displays a characteristic ranking, fairly consistent from test to test.

Before moving on to a more disaggregated examination of these ranking data, let us emphasize once more that these comparative analyses are quite insensitive to absolute progress and reflect only gross differences in the outcomes of progress. The consistent position of the Fort Walton Beach site at the bottom of the academic rankings does not imply, for example, that the children there are making no progress. A comparative glance at the mean WRAT scores in Figures MIV-11b and MIV-11c reveals, in fact, that Fort Walton's mean score gains from fall to spring in kindergarten are quite comparable to those in other sites, but they are not sufficient to close the wide initial gap. Whether Fort Walton's WRAT reading gain from 9.00 to 15.0 represents the same kind of accomplishment as does Seattle's gain from 14.29 to 20.8, on the other hand, is another question with which we cannot deal here.

Portageville's large kindergarten rank gain, on the other hand, does not by itself imply that the Kansas program there is causing greater progress than other programs in other sites are bringing about. Other data on this site, presented elsewhere in this Follow Through report, suggest that the non-Follow Through comparison children in Portageville may also have shown unusual progress, so that the most impressive and obvious features of our results may stem more from site effects than from treatments. Without non-treatment comparison data, we cannot draw any definitive inferences about the role of treatment on bringing about the changes we observe.

The ranking trends shown in Figure MIV-13 represent the aggregation of several dimensions of academic achievement at each testing point. It is informative to consider these dimensions individually, since sponsors intend to emphasize different dimensions in their instruction technique.

The disaggregated longitudinal ranking patterns of Figures MIV-14a through MIV-14c reveal that some substantial changes in rank do indeed occur during Head Start and first grade, for some Sponsors and some dimensions. Portageville's small gain of a single rank in Head Start, for example, turns out to be made up of a large gain (4 ranks) in "reading", a small gain (1 rank) in "math", and a small loss (1 rank) in "vocabulary".

Before we proceed further to interpret these diagrams, the reader is entitled to some explanation of the character of the longitudinal dimensions that they trace. The dimensions of academic achievement vary, of course, among the four testing occasions. During kindergarten and first grade, they comprise mainly the usual reading, mathematics, and vocabulary subject areas, and so we examine the ranks of site means on the corresponding subtests in the kindergarten WRAT and first grade MAT. To trace these dimensions back into the pre-school HSPV experience, where achievement is measured appropriately in terms of more rudimentary skills and knowledge, we must rely upon rather less precise longitudinal analogies.

For the purposes of these traces, we consider the "Numerical" and "Associate Vocabulary" subtests of the Head Start Pre-School Inventory as pretests, at a crude level, for the later WRAT and MAT mathematics and vocabulary subtests, respectively. We take this shaky step the more confidently in the knowledge that the sites fall in very nearly the same rank order with respect to each of the seven pretests tabulated in Figure MIV-11a, whatever the domain being tapped.

For the "reading" dimension, we use as a set of pretest ranks the aggregated achievement pretest ranking used in Figure MIV-13, since the very notion of a "reading" test in preschool has little meaning. A glance across the three pages of Figure MIV-14 will reassure the reader that the

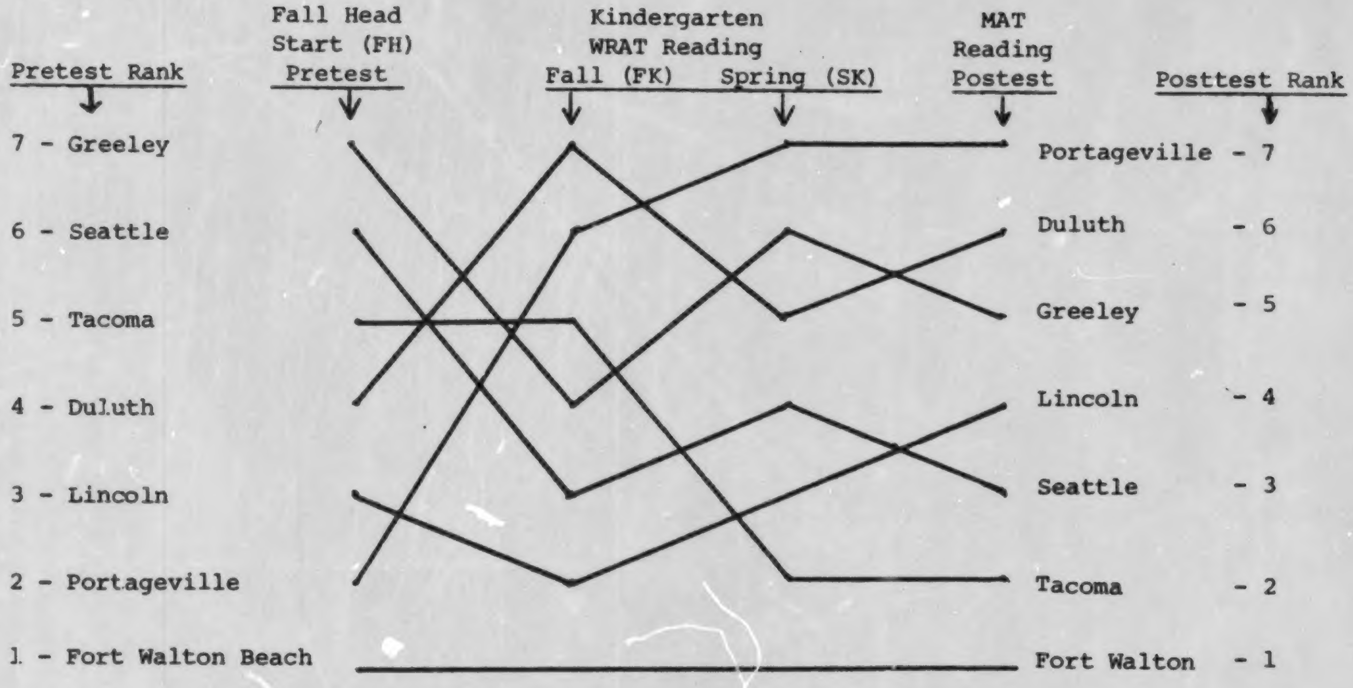


Figure MIV-14a Longitudinal Patterns of Mean Site Ranks of Reading Scores over Four Testing Occasions.

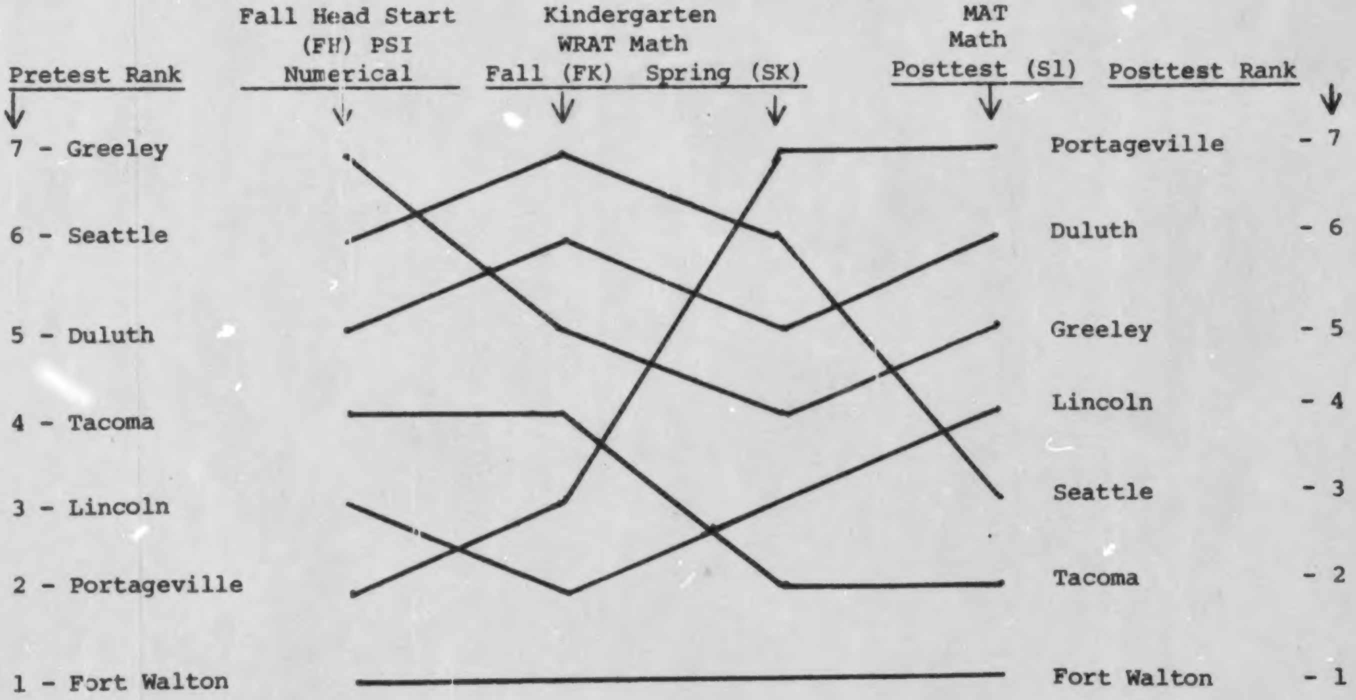


Figure MIV-14b Longitudinal Patterns of Ranks of Site Mean Math Scores over Four Testing Occasions.

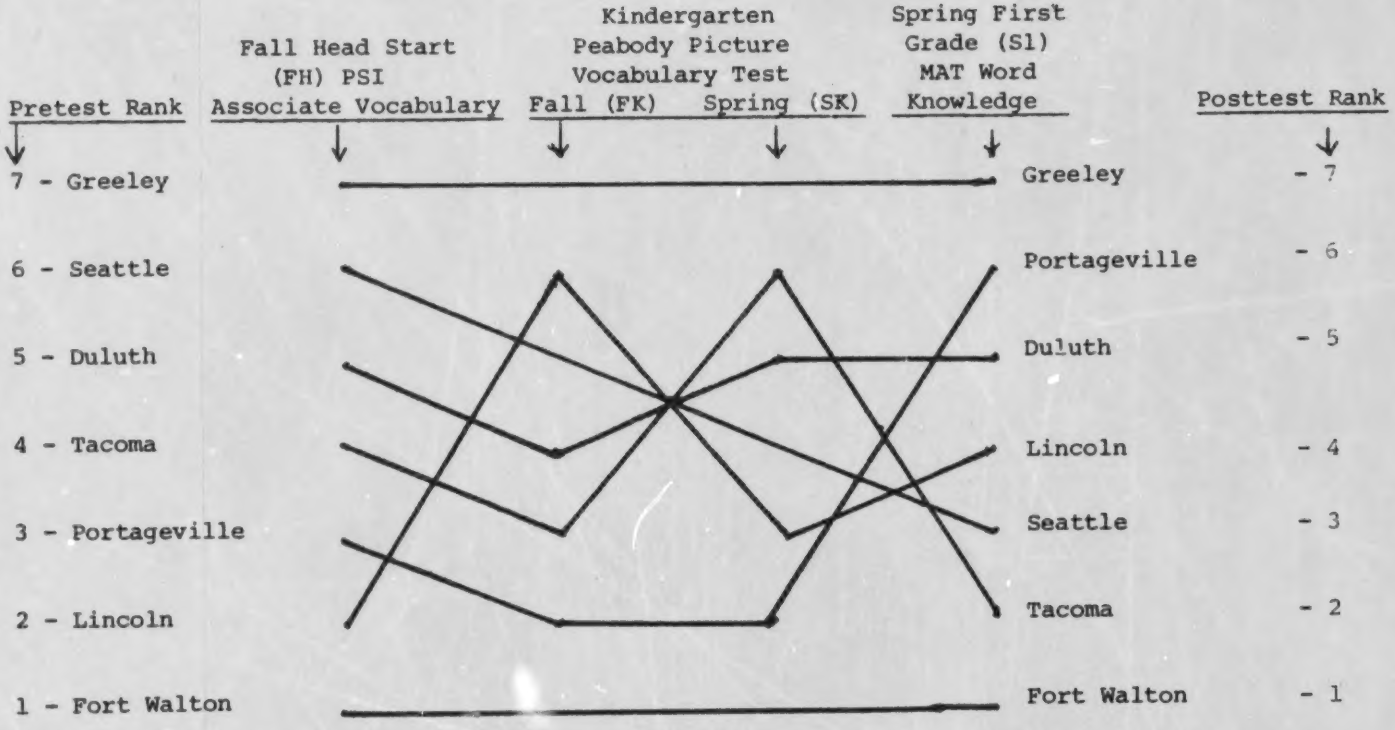


Figure MIV-14c Longitudinal Patterns of Mean Site Ranks of Vocabulary Scores over Four Testing Occasions.

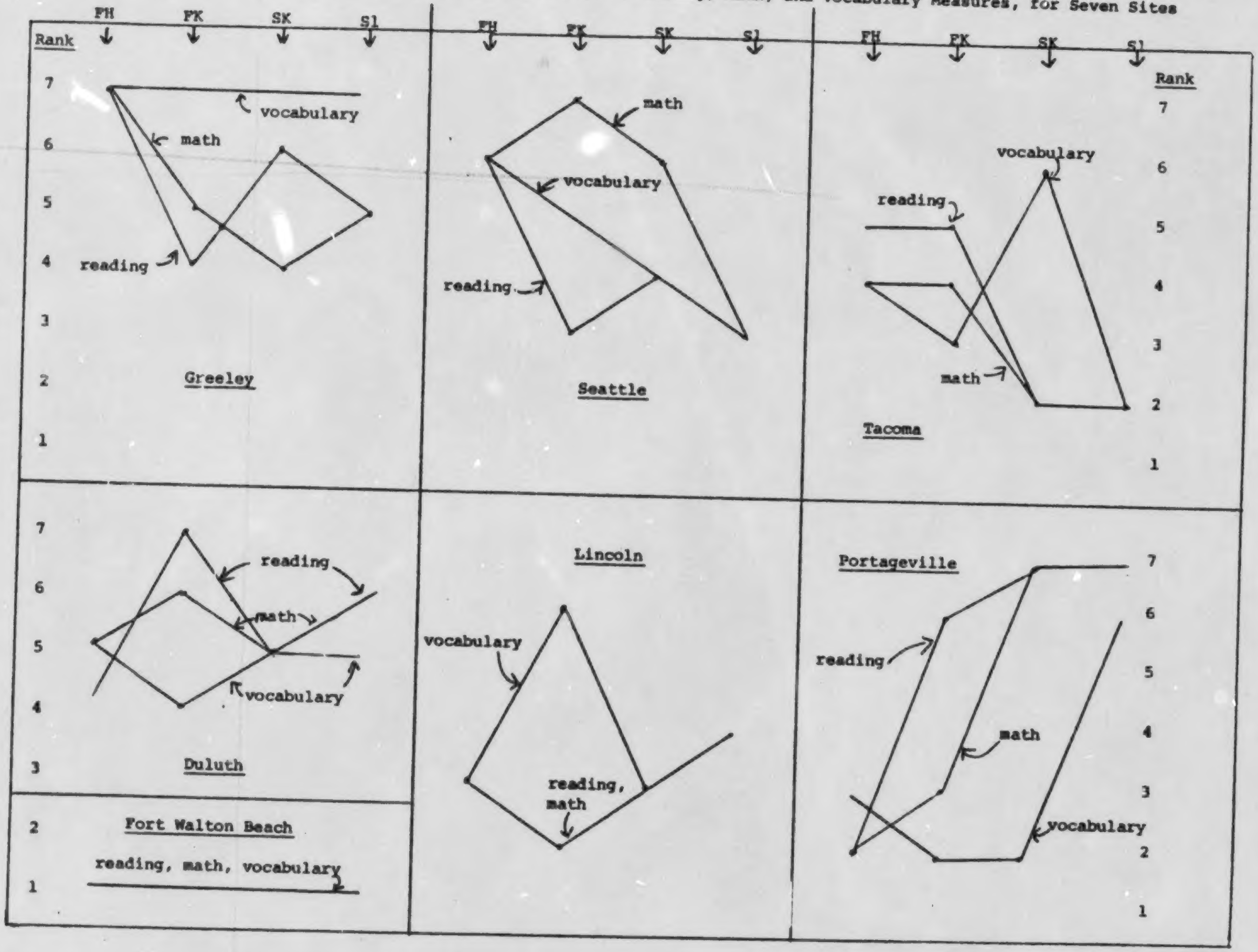
pretest rankings on the three dimensions are indeed very nearly identical--as are, for that matter, the posttest rankings. A single inversion distinguishes the math and vocabulary pretest rankings from the aggregate achievement pretest rankings, and only Greeley's two-rank jump to the top of the vocabulary posttest ranking differentiates among the three posttest rankings.

Within the limits of a study of rankings, it is an interesting finding that each Sponsor seems to have achieved, by the end of first grade, a roughly balanced set of gains with respect to the three principal dimensions of academic achievement. Putting it negatively, no Sponsor seems to be a "reading Sponsor" who neglects math, nor a "vocabulary Sponsor" who leans on rote memory achievements to the exclusion of more conceptual development. Each Sponsor's position in the posttest rankings of overall achievement and of each measured dimension of achievement appears to depend primarily on generalized entry level, and on site-specific factors, as would not be the case if Sponsors "overspecialized".

Given that each Sponsor starts out with a characteristic entry rank, and arrives after three years at a characteristic posttest rank, the question remains whether the three dimensions always proceed in parallel fashion in each Sponsor's hands, or whether some Sponsors achieve early gains on one dimension, leaving the others to catch up later, perhaps with the help of the skills gained during the early thrust. The graphs of Figure MIV-14 provide the answer to this question of dimensional sequence, but the patterns are clearer in Figure MIV-15, where each Sponsor's math, reading, and vocabulary ranking patterns are separated out from those of other Sponsors.

Of all the patterns, Portageville's is the clearest example of what we may call a "phased-progress" pattern. In comparison with the other sites, Portageville made its great leap forward along the reading dimension during Head Start, remaining in the lower ranks of math and vocabulary performance. During kindergarten, while maintaining and even slightly improving its reading rank, Portageville moved to the top in the math rankings, while vocabulary performance remained low. Finally, in first grade, the vocabulary rank increased dramatically to join the other two.

Figure MIV-15 Longitudinal Patterns of Site Mean Scores on Reading, Math, and Vocabulary Measures, for Seven Sites



Duluth shows a similar phased-progress pattern, although rather less clearly, since its overall pre-post rank gains are quite small. Duluth shows, however, the same reading-math-vocabulary sequence as does Portageville.

Since the set of ranks has a fixed total, the rank gains of Portageville necessarily imply corresponding losses in rank by other sites. Its Head Start reading gains, for example, seem to come at the expense of compensating relative losses in Greeley and Seattle; its kindergarten math gains from Greeley and Tacoma; and its first grade vocabulary gains mainly from Tacoma. The very clarity of Portageville's pattern thus tends to make it more difficult to interpret the other patterns.

Figure MIV-16 summarizes the gains in rank for each dimension during each year. In addition to Portageville's three major gains, we note also that Duluth and Lincoln gained substantially during Head Start, in reading and vocabulary respectively, and that Tacoma made a sizable kindergarten vocabulary rank gain. Since a ranking procedure effectively guarantees that some gains and some losses will emerge, it is hard to make much of the smaller rank shifts except, possibly, in the neighborhood of the larger ones we have just noted. With Portageville acting to raise the average in kindergarten and first-grade vocabulary, Lincoln's single-rank gains (in these cells) may indeed reflect a significant accomplishment. Similarly, Duluth competes successfully with Tacoma for position in the kindergarten vocabulary rankings.

No site achieved a major surge in math scores during Head Start, relative to the other sites. Only Lincoln gained rapidly in vocabulary during Head Start, with unimpressive later results in the other dimensions. Wherever reading lost ground at first, however, the overall trend was downward (except, once again, in Lincoln, where it was ambiguous).

These findings obviously cannot support prescriptive conclusions, especially since they embody only comparisons among Sponsors' treatment groups, with no reference to "untreated" NFT or control groups. The FT/NFT contrasts reported elsewhere in this Follow Through report present rather a different picture of Portageville's accomplishments, for example, compared to its own local controls, Portageville lost considerable ground during first grade in its reading scores, suggesting that some site-specific factor may have enhanced both FT and NFT reading progress in Portageville, making them both look good compared to FT and NFT in other sites but unimpressive compared to each other.

Progress Dimension	Year in HSPV-FT		
	Head Start	Kindergarten	First Grade
Reading	Portageville* Duluth*	Greeley Seattle Lincoln Portageville	Duluth Lincoln
Math	Portageville Seattle Duluth	Portageville* Lincoln	Greeley Duluth Lincoln
Vocabulary	Lincoln*	Tacoma* Duluth	Portageville* Lincoln

6

*Rank change ≥ 3 ranks

Figure MIV-16 Sites Showing Gains in Rank Standing on Three Dimensions of Progress During Three Years in HSPV-FT

5.0 Conclusions

As we have noted here and there throughout this monograph, we can really conclude from these data only that we wish heartily that certain additional bits of data had survived to the point of analysis. In the absence of usable data on non-Planned Variation (NPV) children in Cohort III, our analyses can neither support nor challenge the hopeful expectation that preschool sponsorship should confer long-term academic benefits on participating children. Nor can we declare some Sponsor's effects long-lasting and others' mere flashes in the pan. In brief, we have no real evidence on the effects of extended sponsorship in Cohort III. Some of the patterns that emerge from our crippled manipulations take suggestive forms, but the most interesting of them do not hold up in the context of the better-controlled Follow Through findings in the same sites.

Unless additional NPV children can be found somehow within the tested samples of Follow Through Cohort III, we must conclude that the policy-maker will derive no useful information from any further HSPV studies in this Cohort. If the 67 NPV cases in Cohort IV are carefully husbanded and receive a full sequence of testing in subsequent grades, it may yet be possible to learn something valuable about the effects of continuous sponsorship, in those sites at least where the comparison children are concentrated, and especially with respect to the affective dimensions which have been measured much more thoroughly in Cohort IV than in Cohort III. Valid Sponsor comparisons from these data are patently out of the question.

APPENDIX A:

SCORING PROCEDURES

APPENDIX A

SCORING PROCEDURES

The PPVT was administered individually according to the publisher's instructions: first, a basal point in the item sequence was found, where the pupil scored eight consecutive correct responses; then administration continued through the sequence of increasingly difficult items until the pupil missed six of eight consecutive items. Since the reliability of the more difficult items is questionable, we scored only the first 75 items for these analyses. Section 4.3 presents our rationale for this decision.

The three subtests of the kindergarten MAT (Listening for Sounds, Reading and Numbers) were scored according to the publisher's instructions.

As originally developed by the publisher, the WRAT includes three subtests: Spelling, Reading and Arithmetic. Each of these was administered in a shortened version; the items used were scored in accordance with the publisher's instructions. Generally, the shortening of the subtests appears to have caused few difficulties. Section 4.3 contains a more detailed discussion of this issue.

Five 12-item subtests and a total score were developed for the Gumpgookies. The subtests were scored according to the author's instructions, using unit-weighted items. The five subtests are Instrumental Activity, School Enjoyment, Self-Evaluation, Purposive Behavior, and Self-Confidence. Because of distributional problems discussed below, we have used only the Instrumental Activity subtest and the total test score in these analyses.

For the Locus of Control Test, internal-positive and internal-negative subtests were scored. The items for Locus of Control-positive measure the the child's acceptance of responsibility for the good events which occur; the items for Locus of Control-negative measure his acceptance of responsibility for bad events.

The subtests of the Spring 1973 MAT were scored according to the publisher's instructions. Word Analysis, Reading, Total Reading and Mathematics were employed in these analyses. For distributional reasons discussed below, Word Knowledge was not used. Because the Total Reading score includes the Reading subtest, the Reading subtest was not analyzed.

The following is a summary of the tests and subtests used in the analyses:

- PPVT, Total
- MAT (Primer, Form F)
  - Listening for Sounds
  - Reading
  - Numbers
- WRAT (Level I)
  - Spelling
  - Reading
  - Arithmetic
- Gumpgookies (Total Score)
  - Instrumental Activity
- Locus of Control
  - -positive
  - -negative
- MAT (Primary I, Form F)
  - Word Analysis
  - Reading total (Reading plus Word Knowledge)
  - Mathematics

APPENDIX B:

TRANSFORMATIONS

APPENDIX B

TRANSFORMATIONS

Since our analyses assume that all outcome measures are normally distributed, we have examined the distribution of each outcome in a sample consisting of NFT children who took the Fall 1971 and Spring 1972 WRAT tests and the Spring MAT. Many of the measures are approximately normally distributed, but some deviate severely enough to warrant transformation or elimination from the analysis.

We have used only monotonic transformations; each transformed score, therefore, retains its original relative position within the distribution. The transformations have no effect on rank correlation coefficients and little effect on the Pearson product-moment correlations. The test-retest reliabilities of our transformed scores may be better (and should be no worse) than those of the original scores. We perform transformations, of course, for the sake of the improved statistical characterisitcs of the transformed scores. Transformed means, for example, are better estimators of population means. Transformed standard deviations and standard errors have more nearly the classical meanings: $\mu \pm 1.96\sigma$ of the transformed variable contains approximately 95 percent of the distribution.

The ill-distributed measures included the two Absence variables, Gumpgookies, Kindergarten MAT Word Knowledge, and First grade MAT Word Analysis.

Figures B-1 and B-3 display the non-normal distributions of Kindergarten Days Absent and First Grade Days Absent. The long right-hand tails and almost nonexistent left-hand tails of these two distributions suggest that a logarithm transformation might be appropriate. This choice is plausible on theoretical grounds; moreover, if each pupil is absent on randomly chosen days, then the distribution of total days absent should approximate a Poisson or log-normal distribution. Since the range of the Absence variables includes zero, which has no logarithm, we have added to each raw absence score before taking the common logarithm. This transformation has little effect on the left-hand tail of each distribution but greatly shrinks the right-hand tail, as Figures B-2 and B-4 clearly show.

The Gumpgookies total score distribution (Figure B-5) has a short left-hand tail and almost no right-hand tail, resembling the distribution of the sample Pearson product-moment correlation when the population correlation is non-zero. This similarity suggested that Fisher's Z-transformation might remedy the non-normality, and so we applied it after first dividing the score by 61, the smallest divisor which ensures that the denominator in the transformation formula will always be non-zero. Figure B-6 displays the results.

Anomalies similar to those of the Gumpgookies raw score also characterize the Gumpgookies Instrumental Activities subtest score distribution. Again, we used Fisher's Z transformation. The result here is less impressive than with the Gumpgookies total score, the right-hand tail is smaller and the distribution is somewhat more symmetric about the transformed mean as Figures B-7 and B-8 show.

Each of the four other Gumpgookie subtests and the first grade MAT Word Knowledge subtest lacked right-hand tails, reflecting distinct ceiling effects (cf. Figures B-9 through B-13). No transformation can create a tail where none exists and so we dropped these test scores from the analysis. The Gumpgookies subtests scores are included in the Gumpgookies total score, however, and the MAT Word Knowledge is included in the MAT Total Reading.

The bimodal relatively flat distribution of the first grade MAT Word Analysis scores lacks inflection in both of its tails and is somewhat similar to the sampling distribution of proportions; we therefore applied an arcsin transformation. The raw score was divided by 41 prior to the transformation in order to make the number to be transformed less than 1. Comparing histograms of the untransformed and transformed scores in Figures B-4 and B-5, we note that the transformed score distribution is somewhat more symmetric about its mean.

The distributional characteristics of each anomalously-distributed test was examined at the population level including FT and NFT children to learn whether the anomalies also obtain in the larger population. Figures B-16 through B-23 display the distrubitons of the untransformed scores for first grade MAT Word Knowledge, Word Analysis and the Gumpgookies total and subtest scores. As may be noted by comparing these histograms with those of the pooled NFT sample which were discussed above, the distributions are quite similar for the NFT sample and the pooled FT/NFT population. This suggests that whatever may be causing the peculiarities in the NFT sample, it is not likely that the source exists within the criteria employed in selecting this analytic sample. A variety of circumstances might create these distributional problems. For example, some of these tests may simply be too easy for these pupils: at the end of kindergarten the children may know the socially desirable responses for the Gumpgookies; at the end of the first grade, they may be thoroughly familiar with the fairly elementary skills demanded on the Word Knowledge subtest.

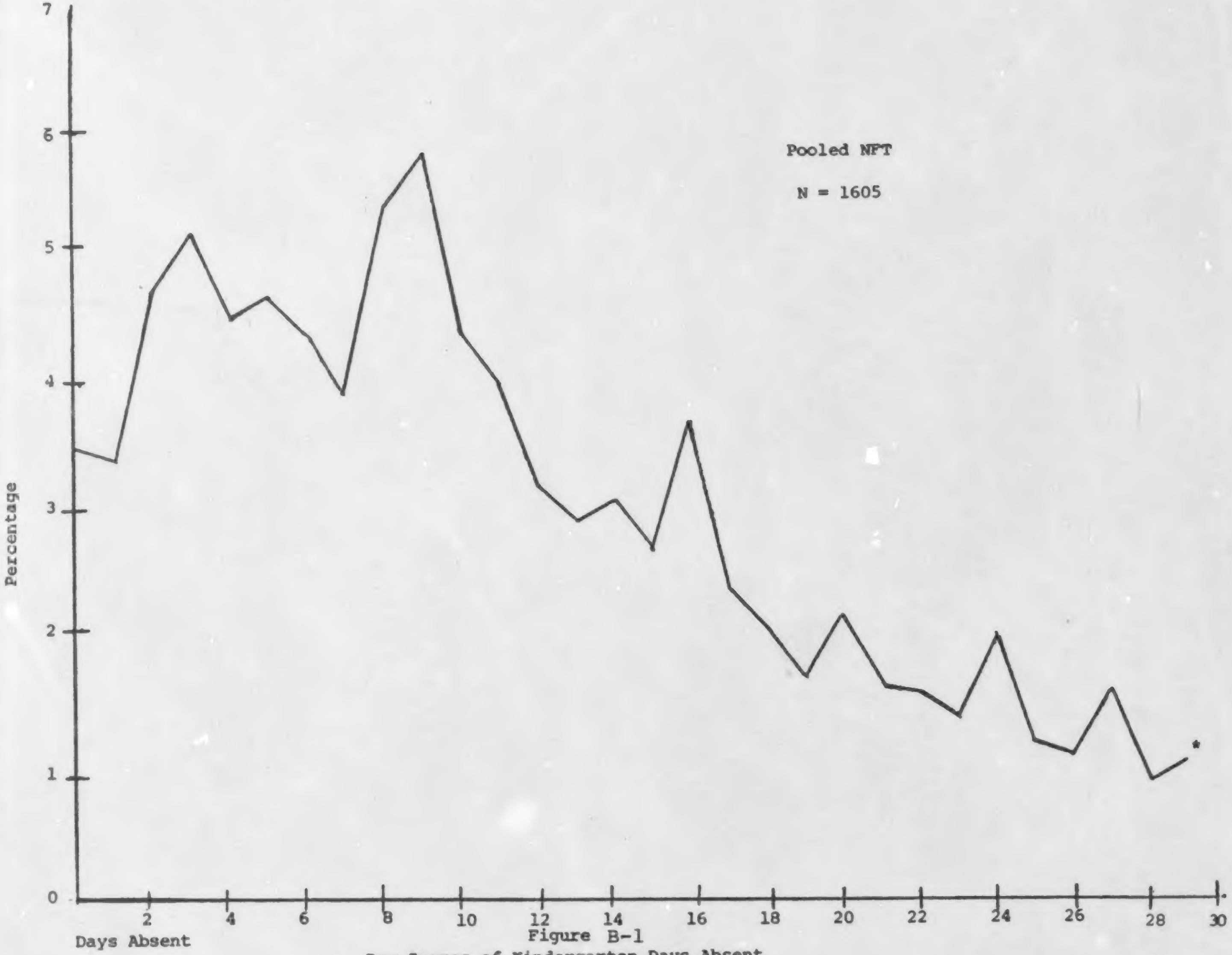


Figure B-1

Raw Scores of Kindergarten Days Absent

*9.8 percent of the children were absent more than 29 days.

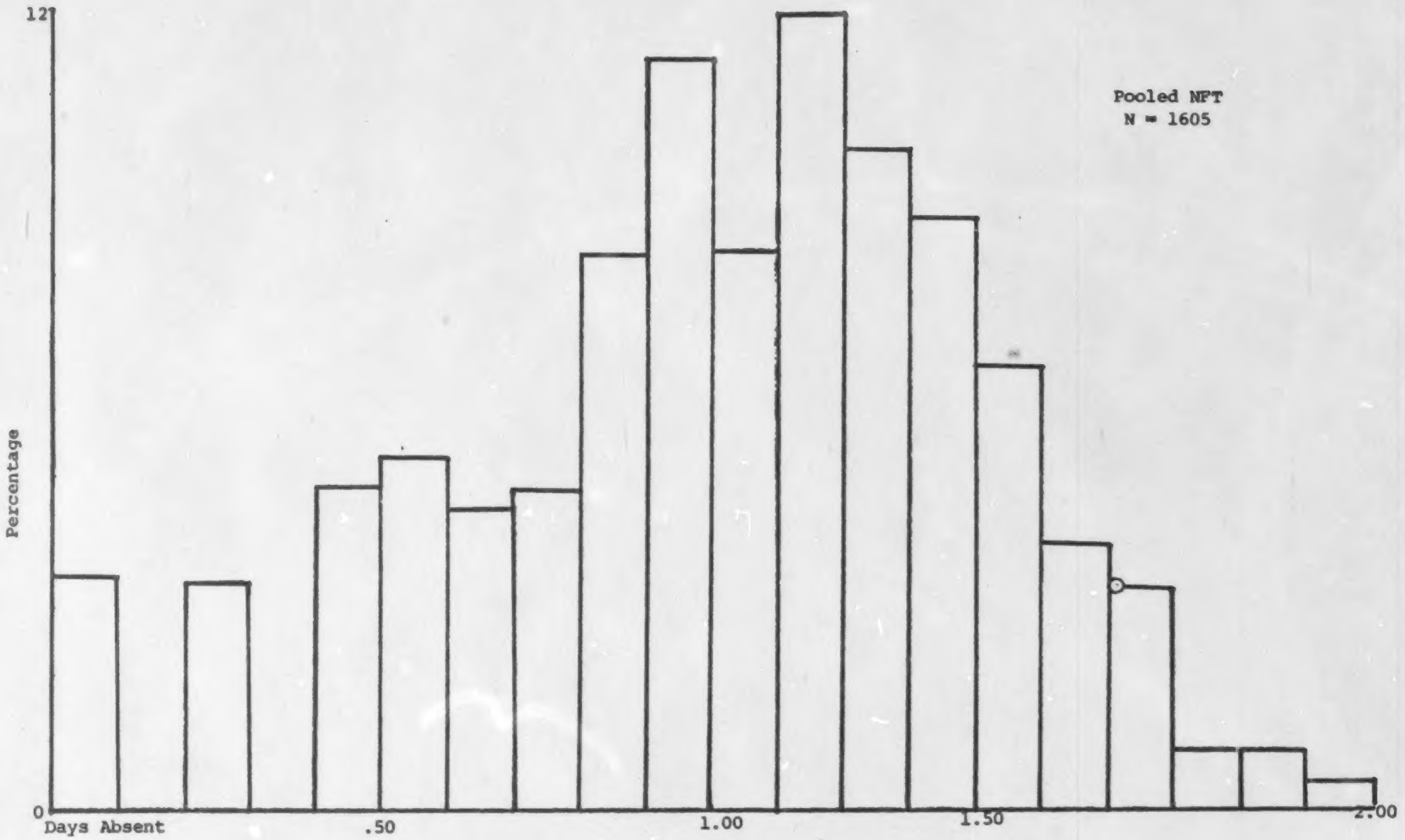


Figure B-2

Transformed Scores of Kindergarten Days Absent
Transformation: Days Absent = $\log_{10}$ (1 + Days Absent)

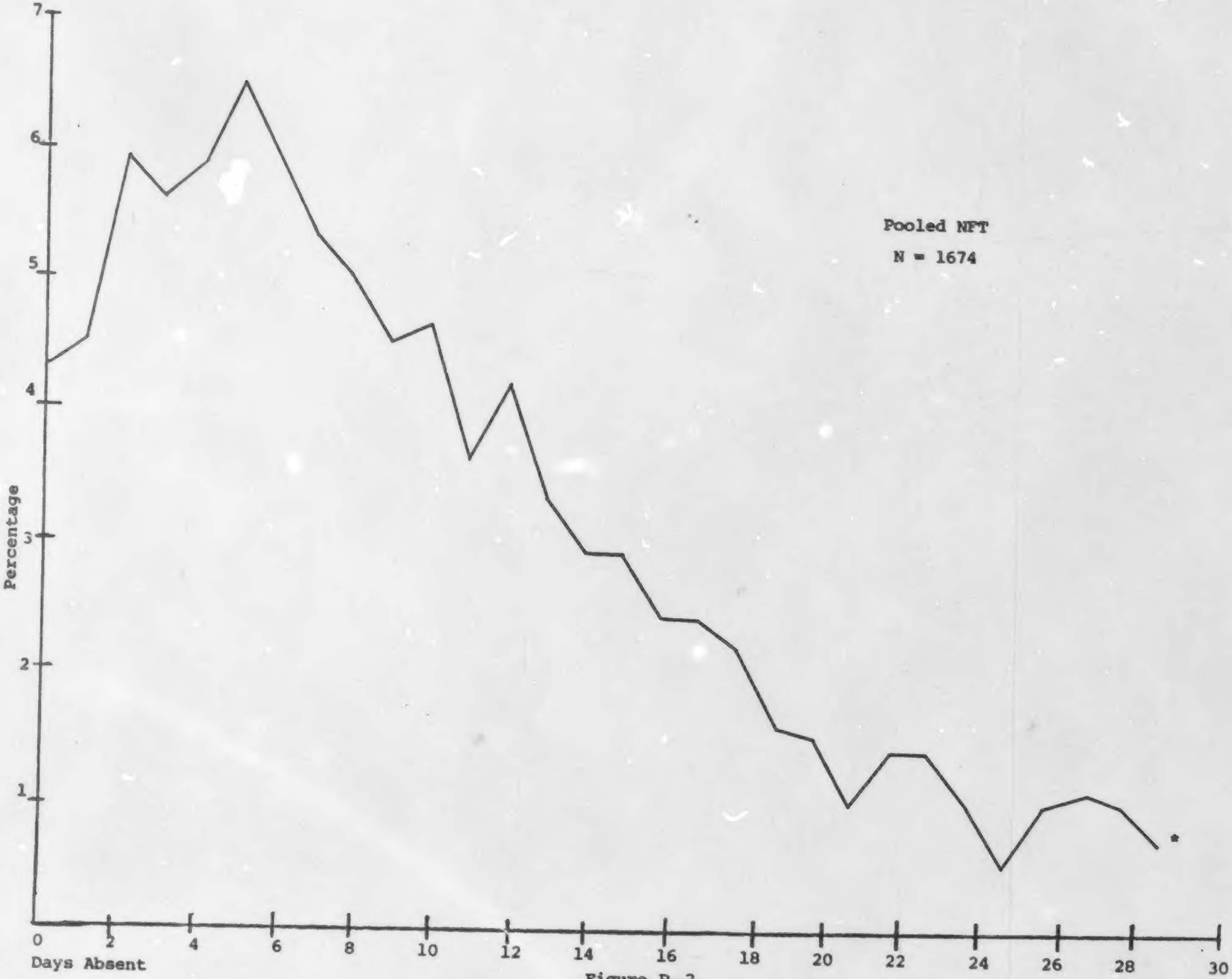


Figure B-3
Raw Scores of First Grade Days Absent

*6.1 percent of the children were absent more than 29 days.

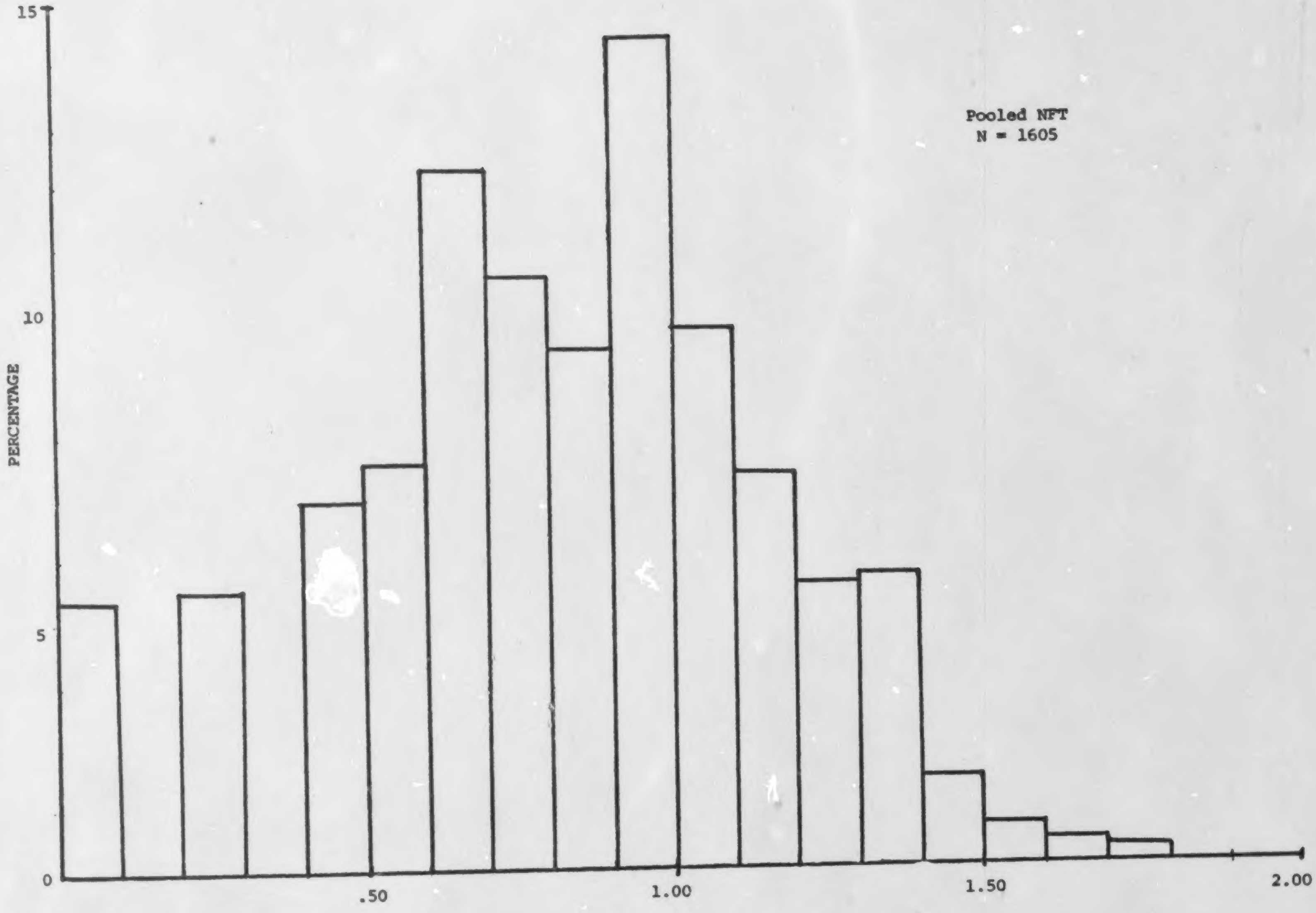
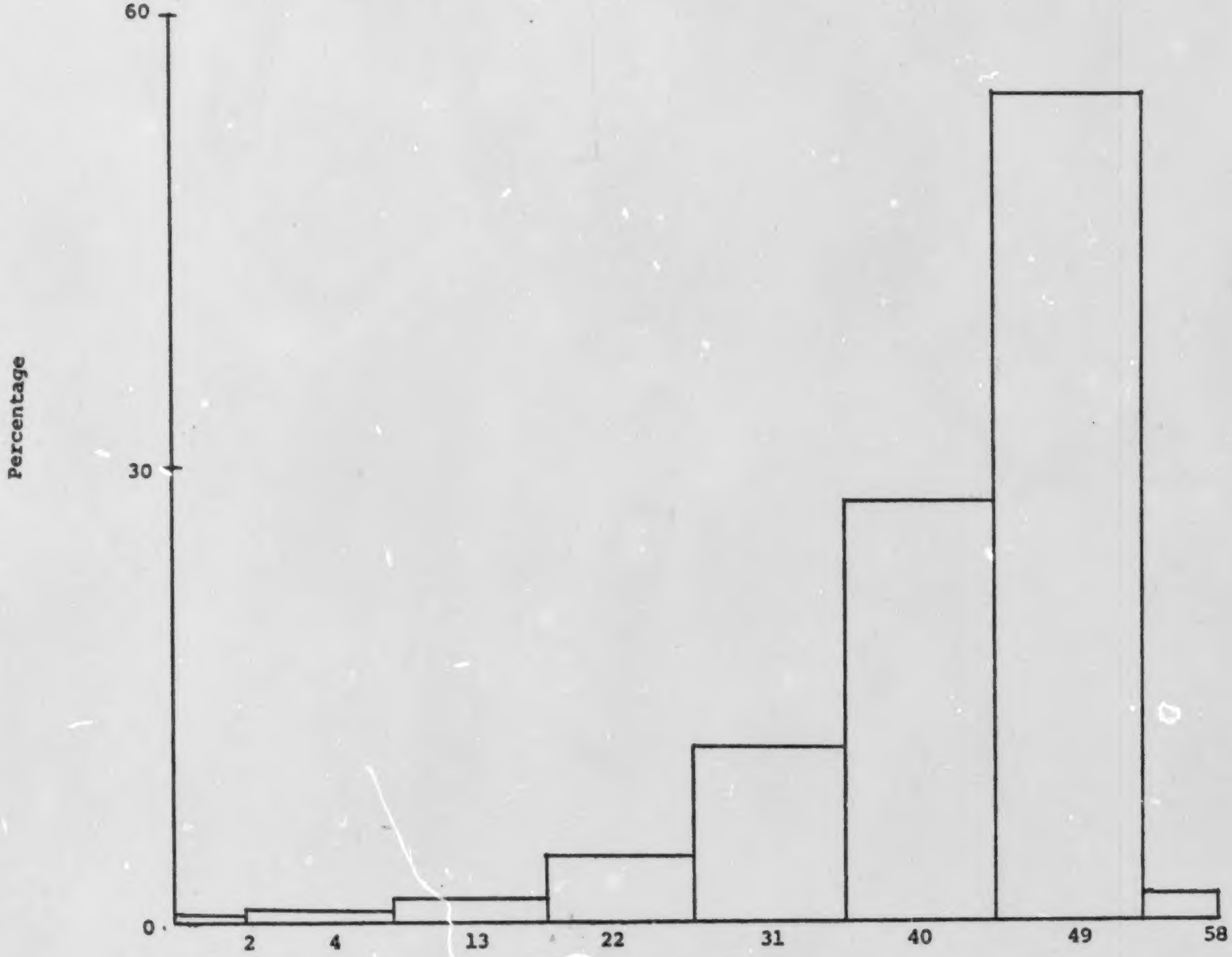


Figure B-4
Transformed Scores of First Grade Days Absent
Transformation: Days Absent = $\log_{10}$ (1 + Days Absent)



POOLED NFT
N = 1630

Figure B-5
Raw Scores of Kindergarten Gumpgookies Total Score

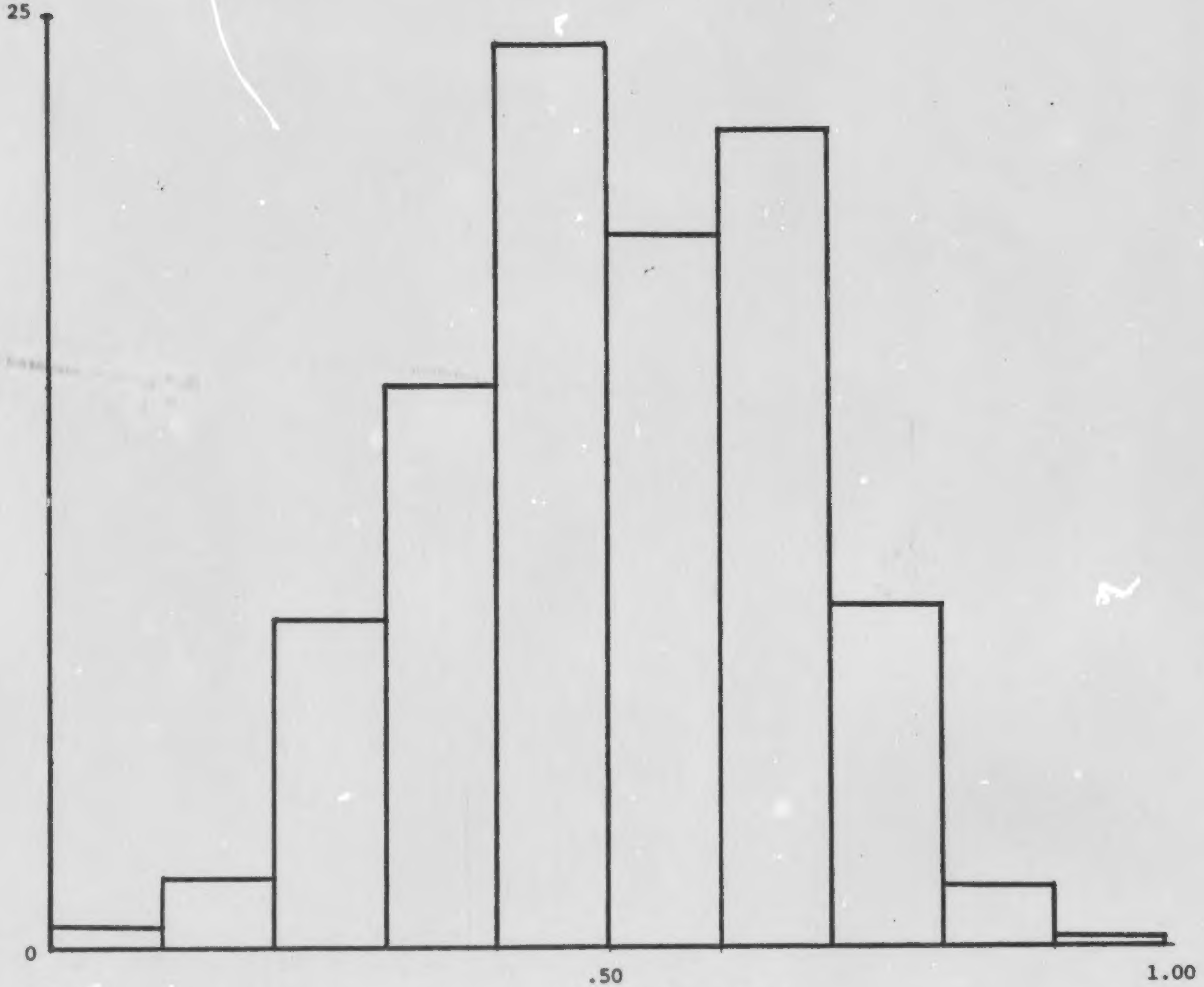


Figure B-6
Transformed Scores of the Kindergarten Gumpgookies Total Score

Transformation: $\frac{1}{2} \log \frac{(1+score/61)}{(1-score/61)}$

Pooled NFT
N = 1630

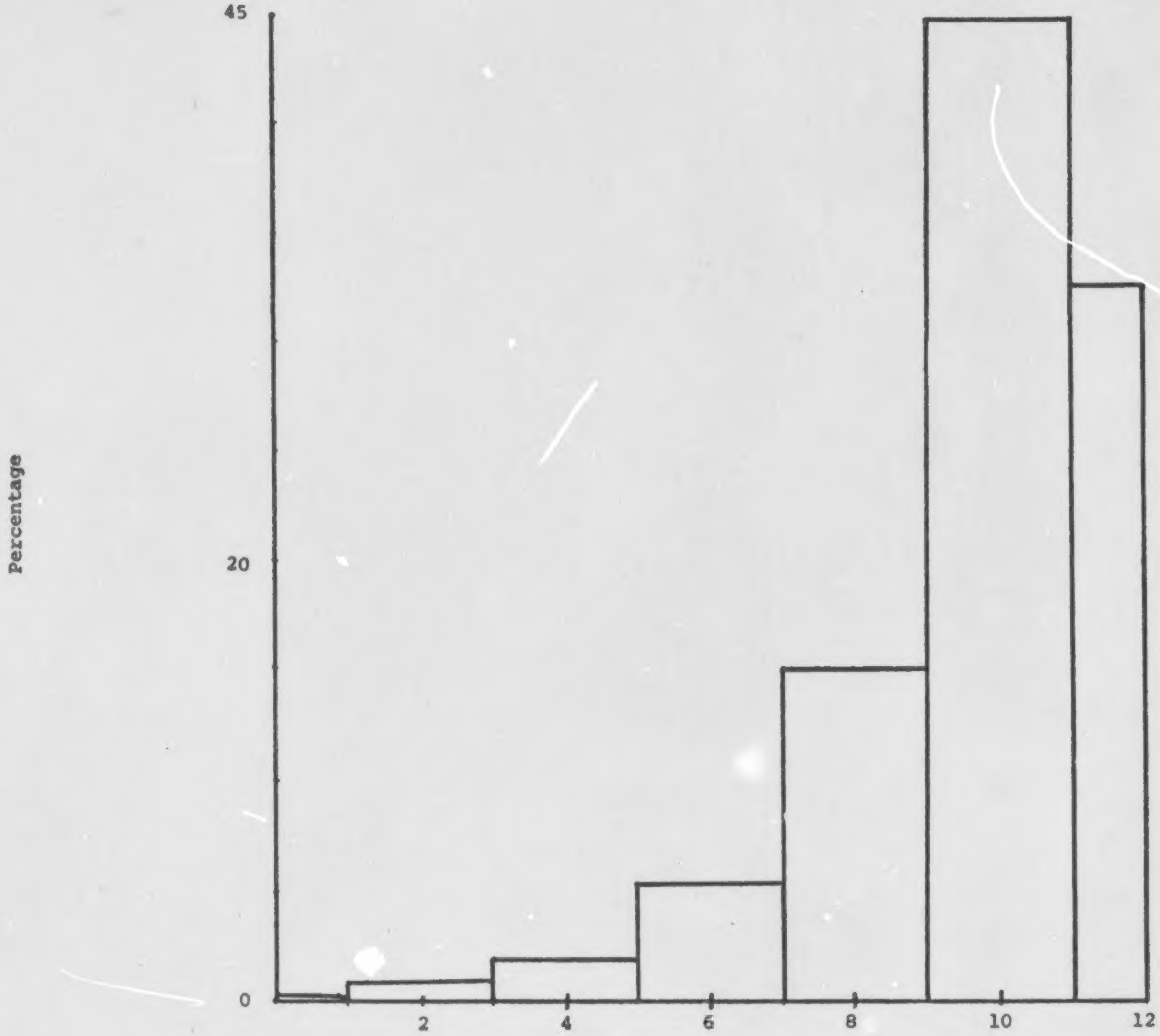


Figure B-7

Raw Score of Kindergarten Gumpgookies Instrumental Activities Subtest

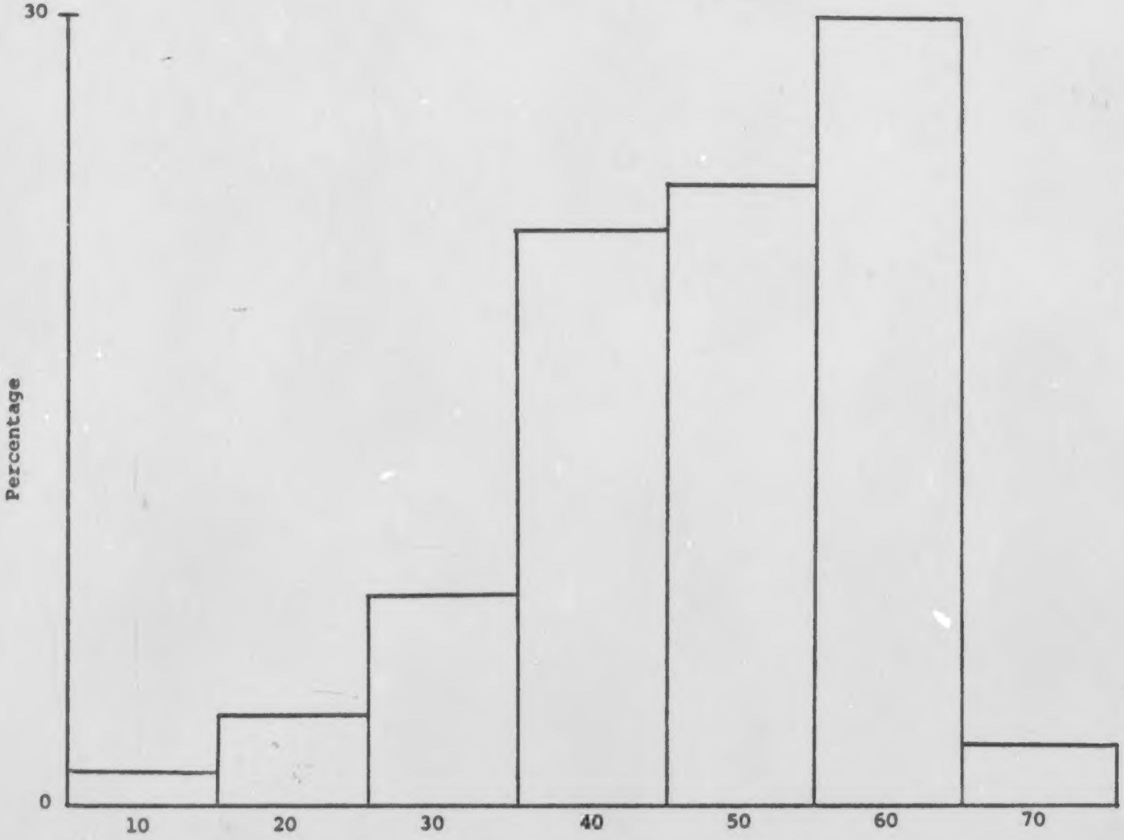


Figure B-8

Transformed Scores of Kindergarten Gumpgookies Instrumental Activities Subtest

Pooled NFT
N=1630

Transformation: $1/2 \log \frac{1+X2}{1-X2}$
X2=Raw Score/13

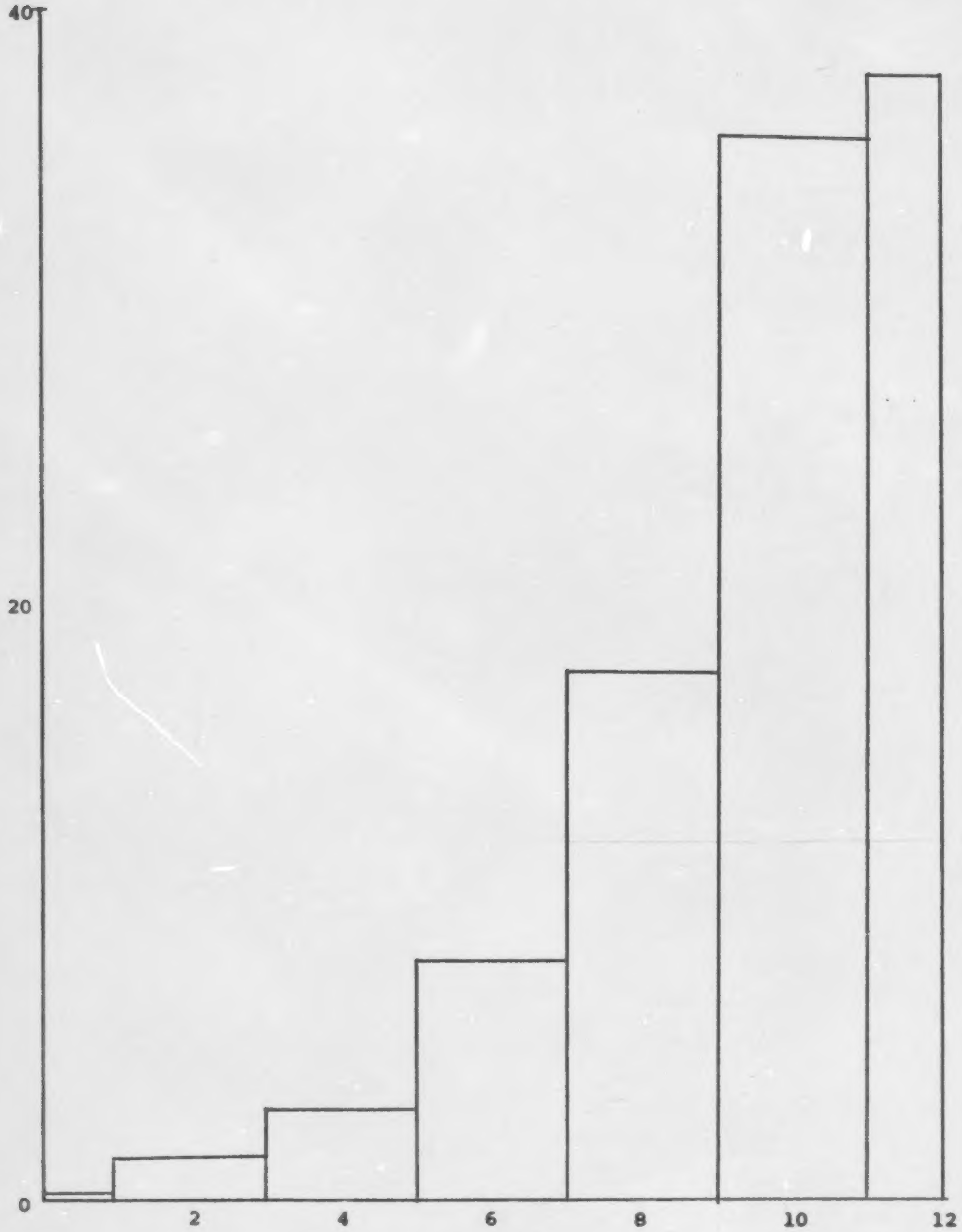


Figure B-9
Raw Scores of Kindergarten Gumpgookies School Enjoyment Subtest

Pooled NFT
N = 1630

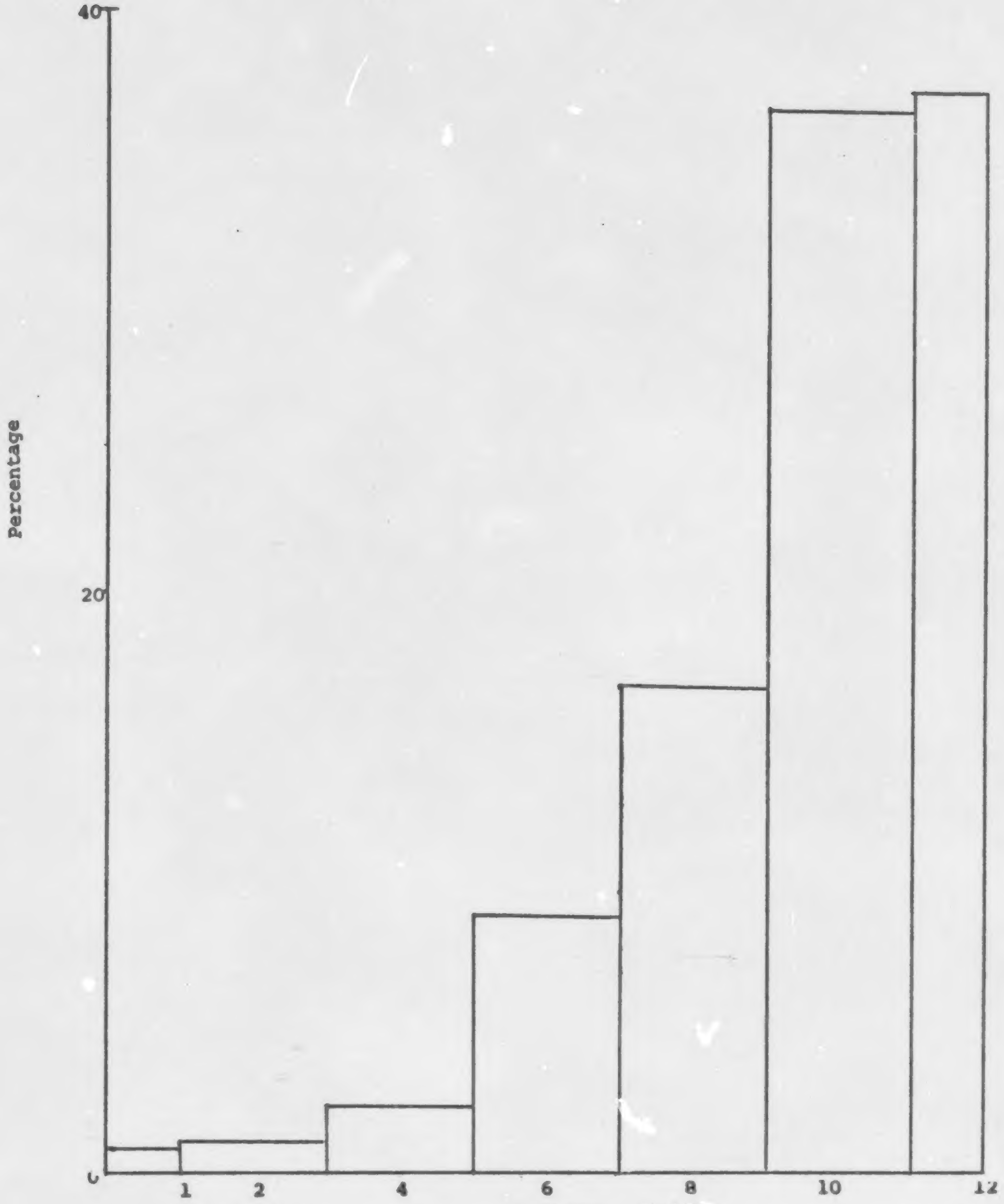


Figure B-10
Raw Scores of Kindergarten Gumpgookies Self-Evaluation Subtest

Pooled NFT
N = 1630

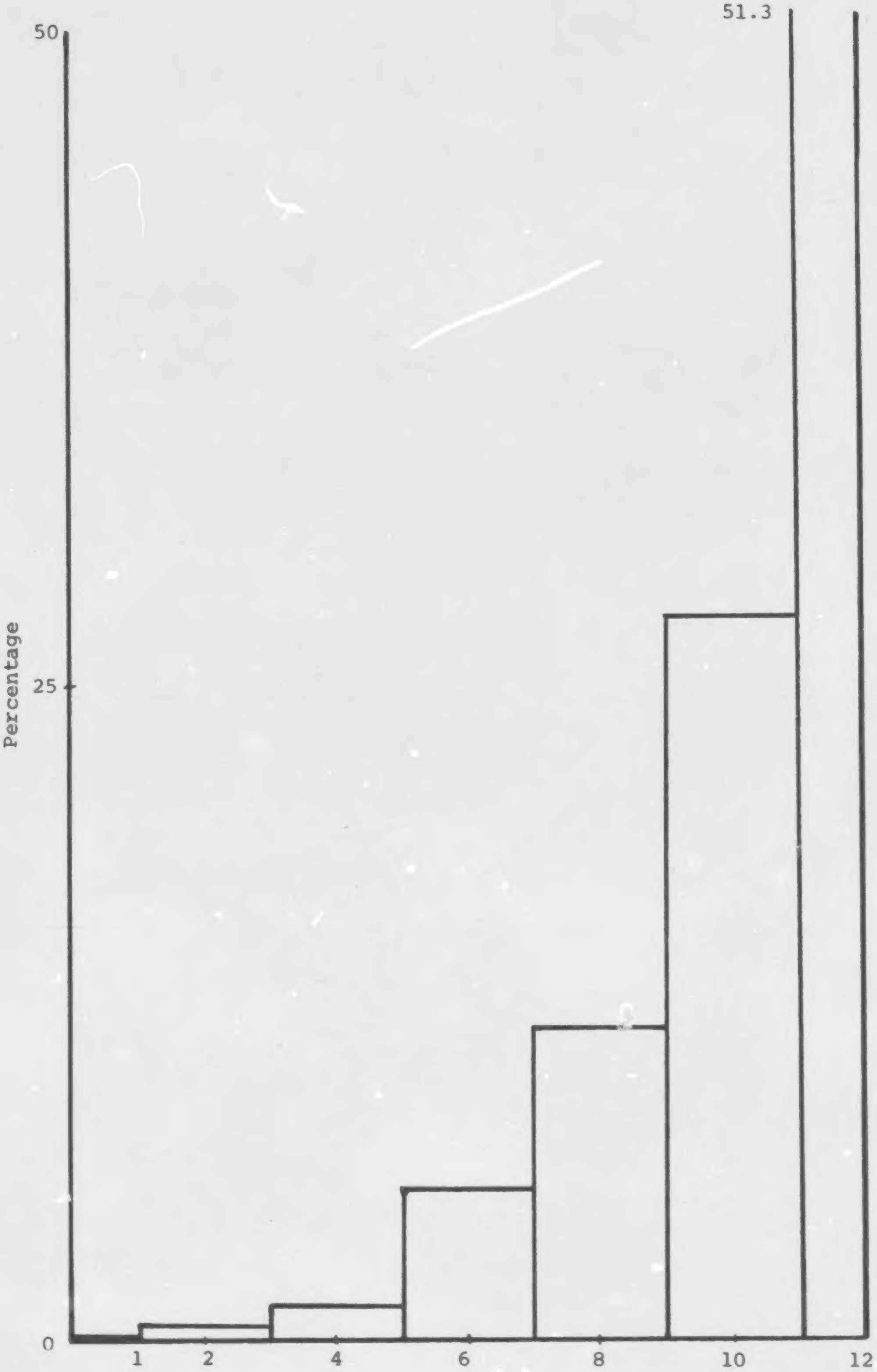


Figure B-11

Raw Scores of Kindergarten Gumpgookies Self-Confidence Subtest

Pooled NFT
N = 1630

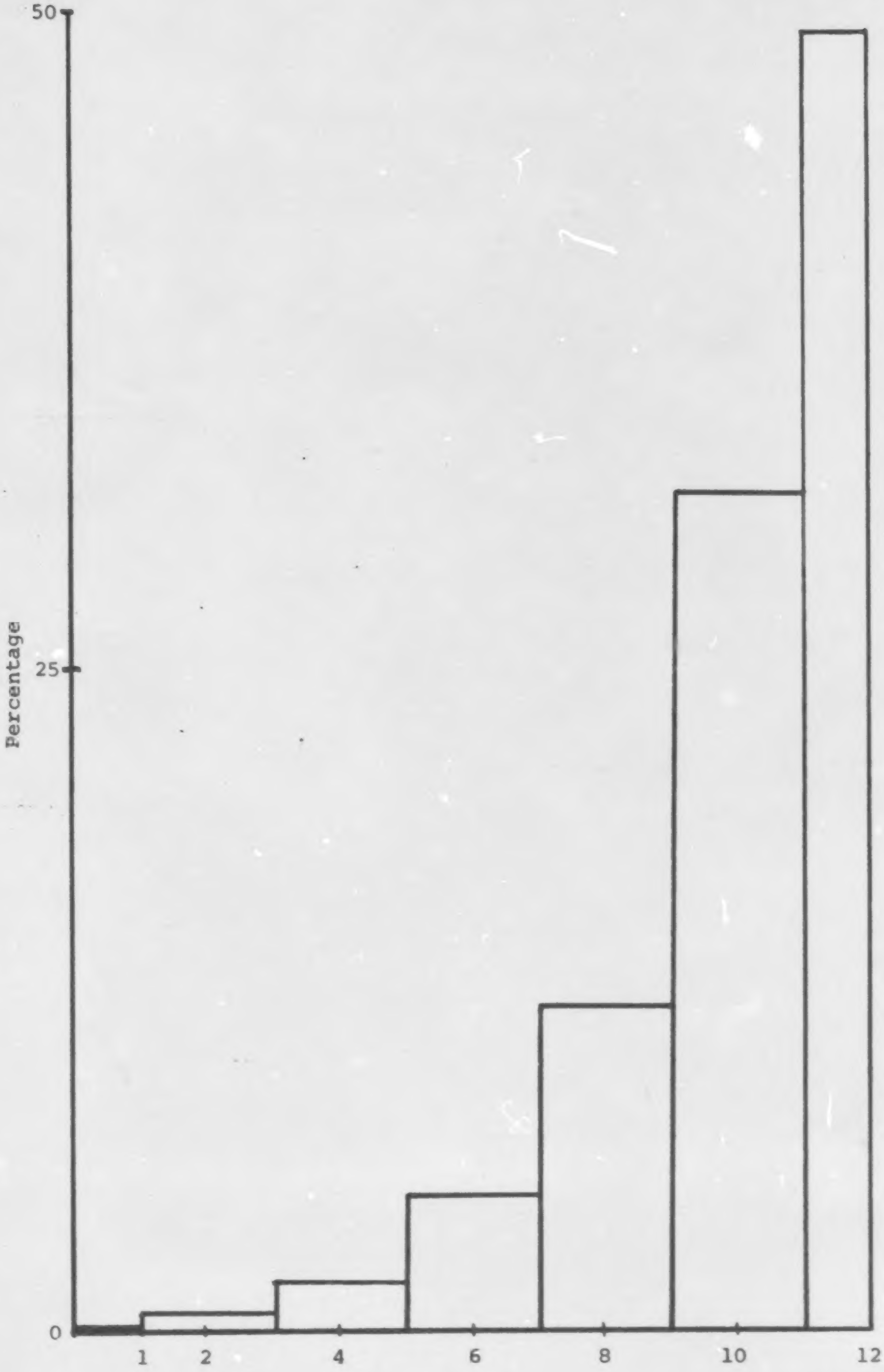


Figure B-12

Raw Scores of Kindergarten Gumpgookies Purposive Behavior Subtest

Pooled NFT
N = 1630

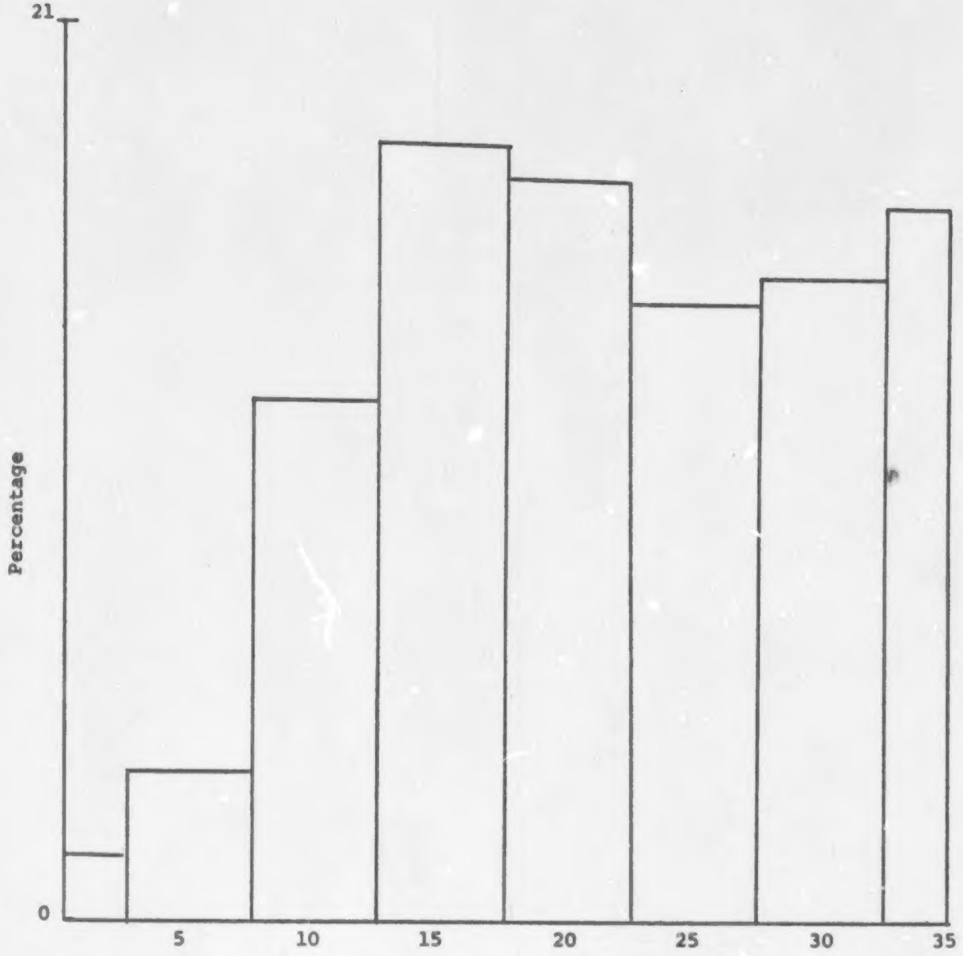


Figure B-13

Raw Scores of First Grade MAT Word Knowledge Subtest

Pooled NFT
N=1683

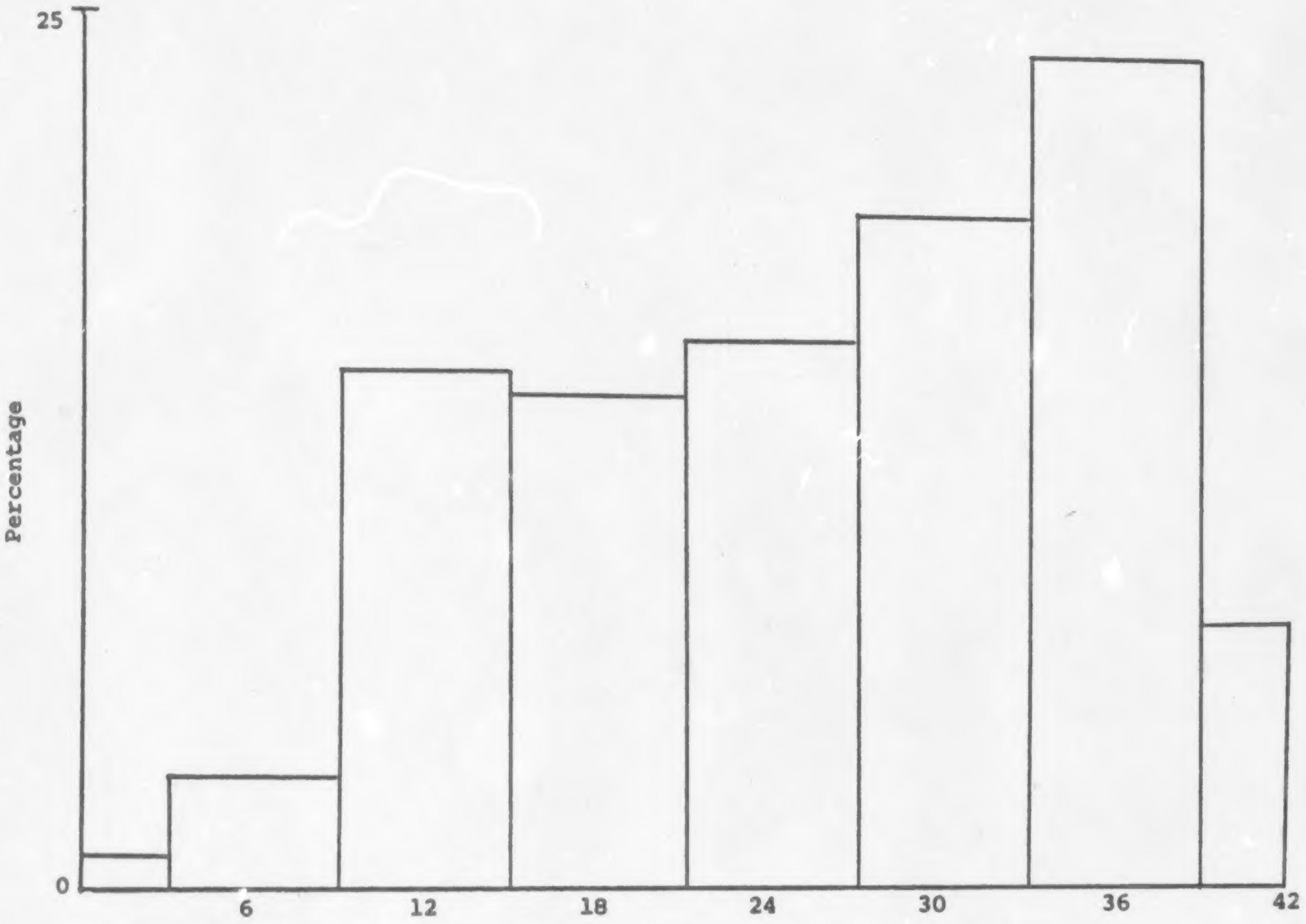


Figure B-14

Raw Scores of First Grade MAT Word Analysis Subtest

Pooled NFT
N=1683

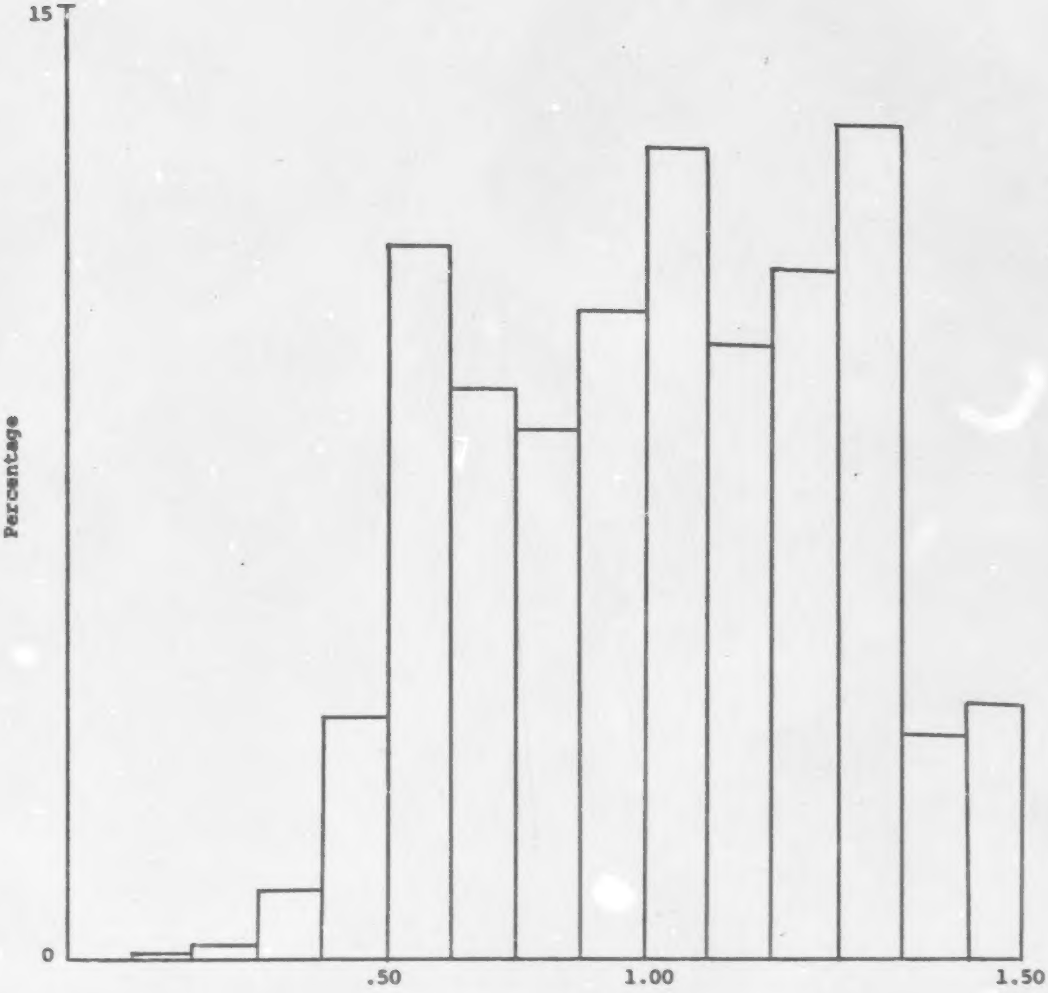


Figure B-15

Transformed Scores of First Grade MAT Word Analysis Subtest

Pooled NFT
N=1683

Transformation: $\arcsin\sqrt{score/41}$

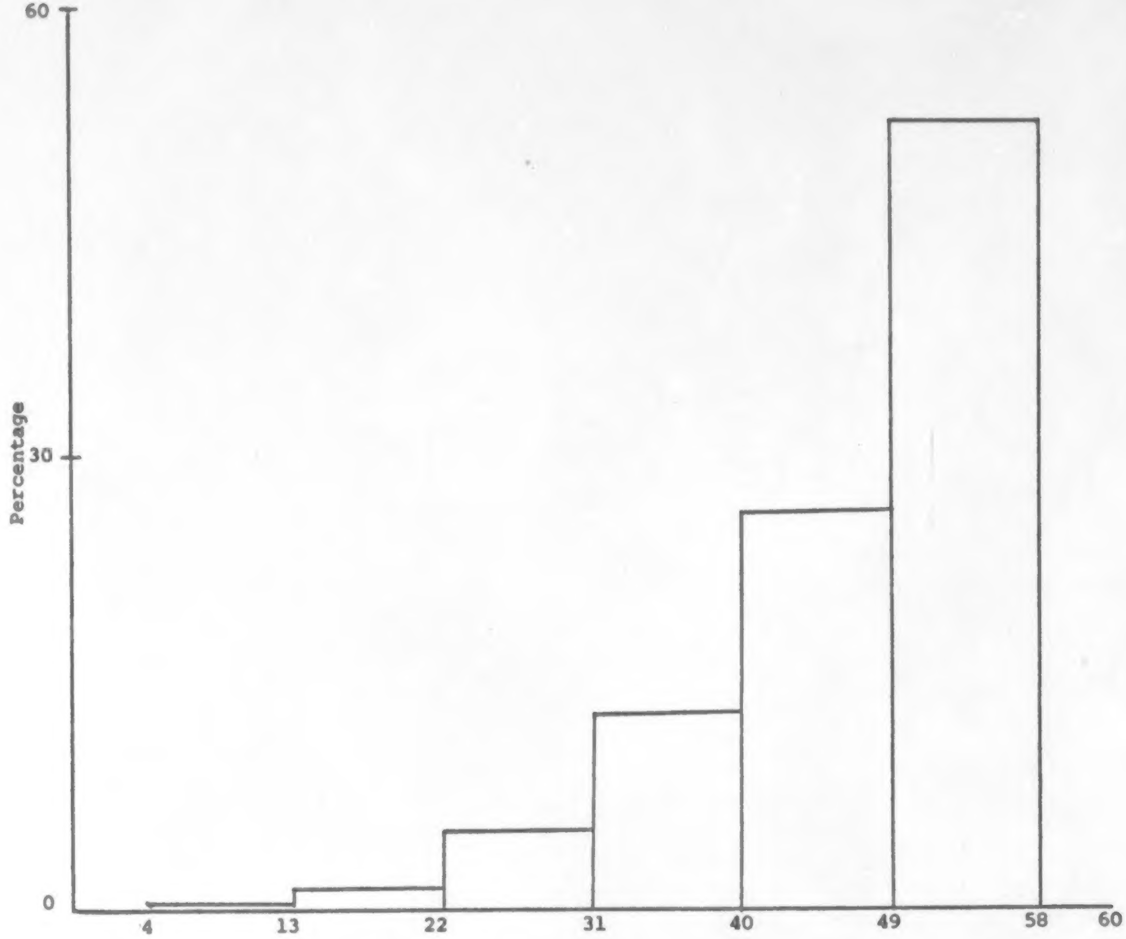


Figure B-16

Raw Scores of Kindergarten Gumpgookies Total Score

Random Sample of Population
N=2688

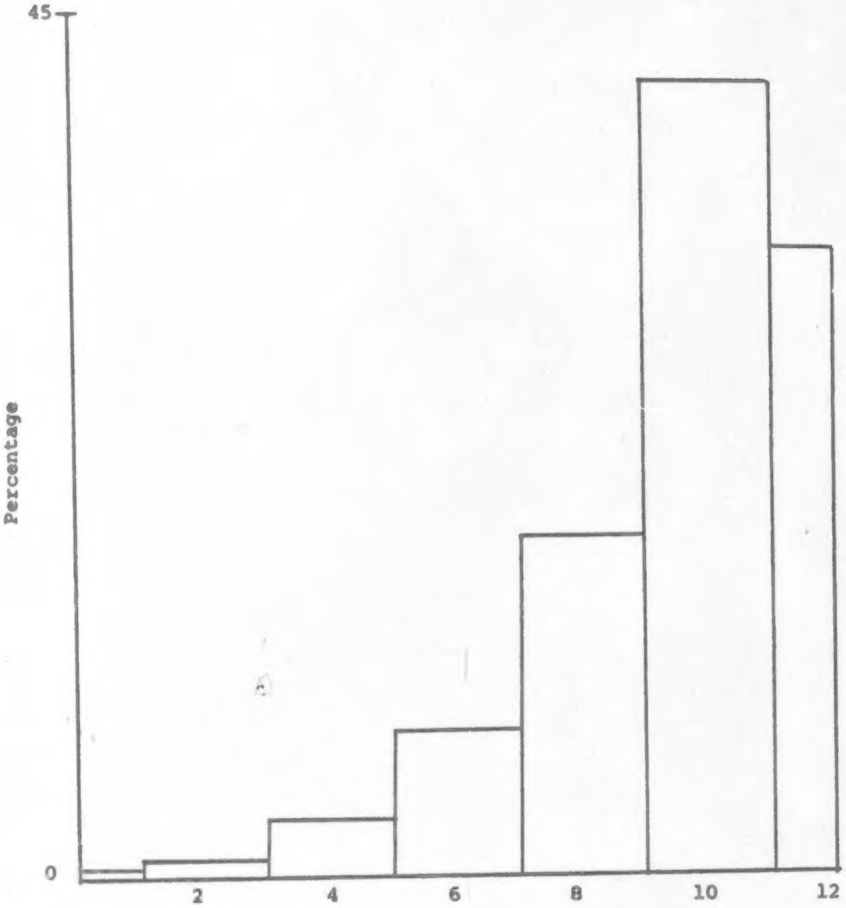


Figure B-17

Raw Scores of Kindergarten Gumpgookies Instrumental Activities Subtest

Random Sample of Population
N=2688

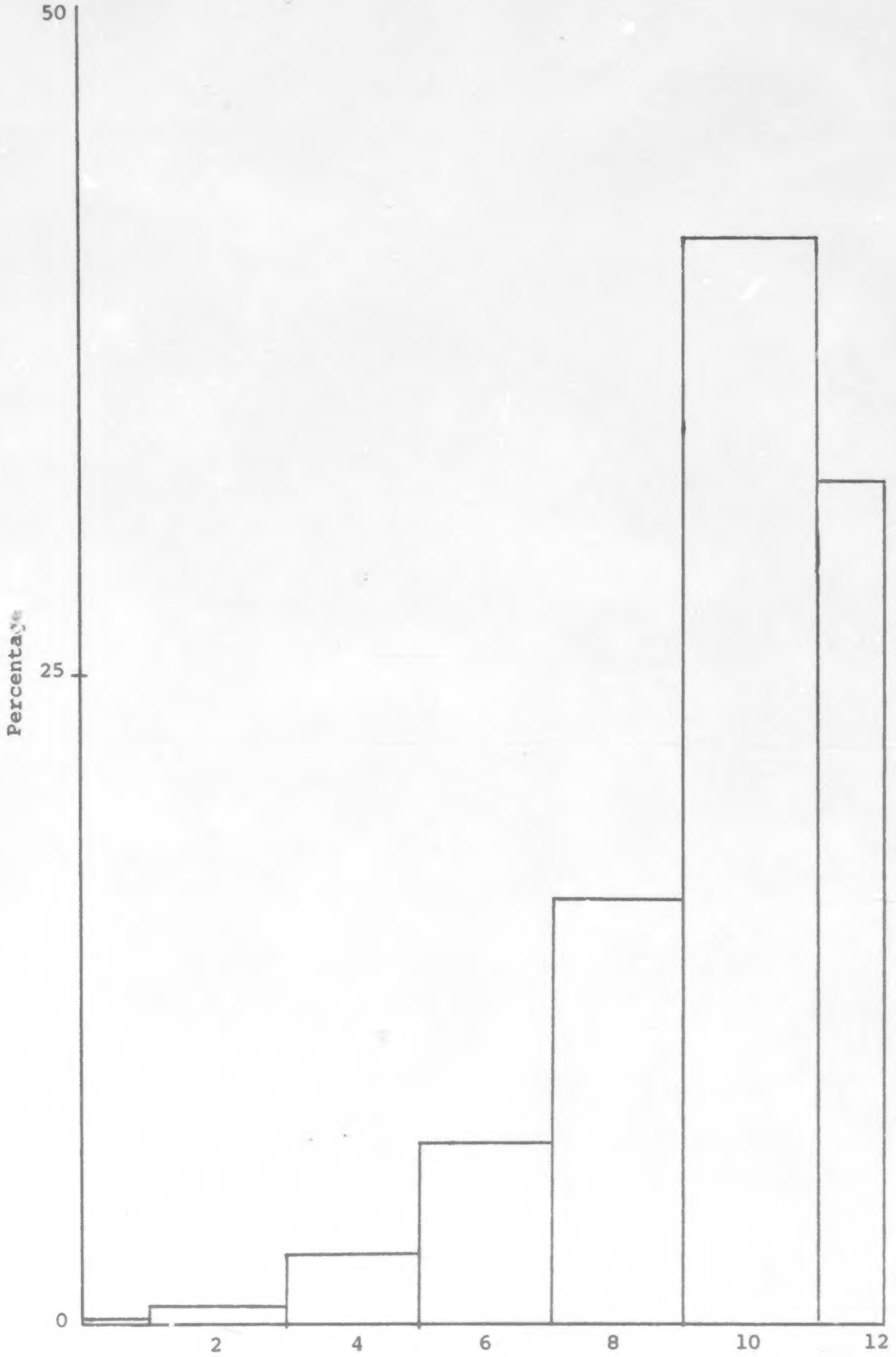


Figure B-18

Raw Scores of Kindergarten Gumpgookies School Enjoyment Subtest

Random Sample of Population
N=2688

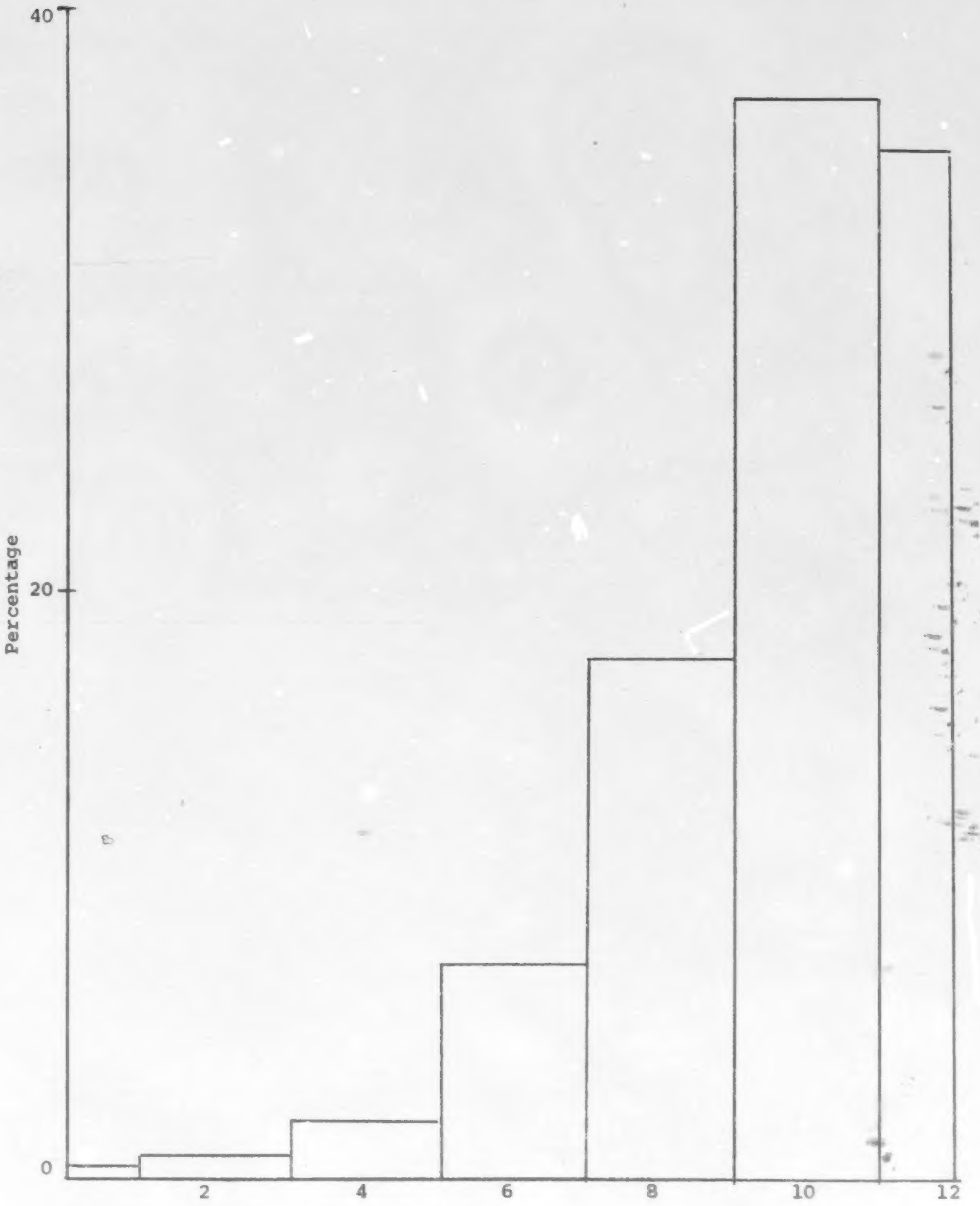


Figure B-19

Raw Scores of Kindergarten Gumpgookies Self-Evaluation Subtest

Random Sample of Population
N=2688

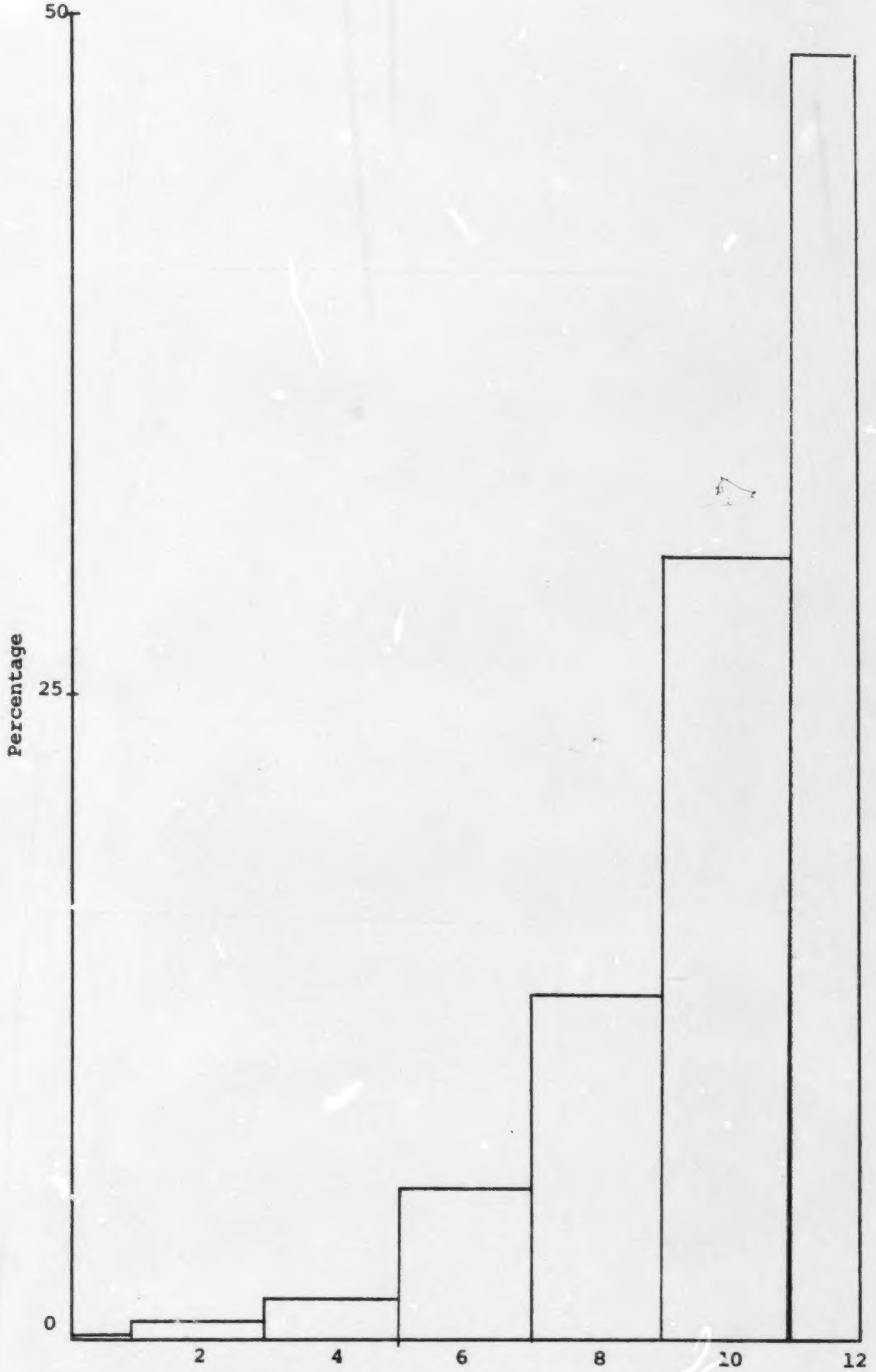


Figure B-20

Raw Scores of Kindergarten Gumpgookies Self-Confidence Subtest

Random Sample of Population
N=2688

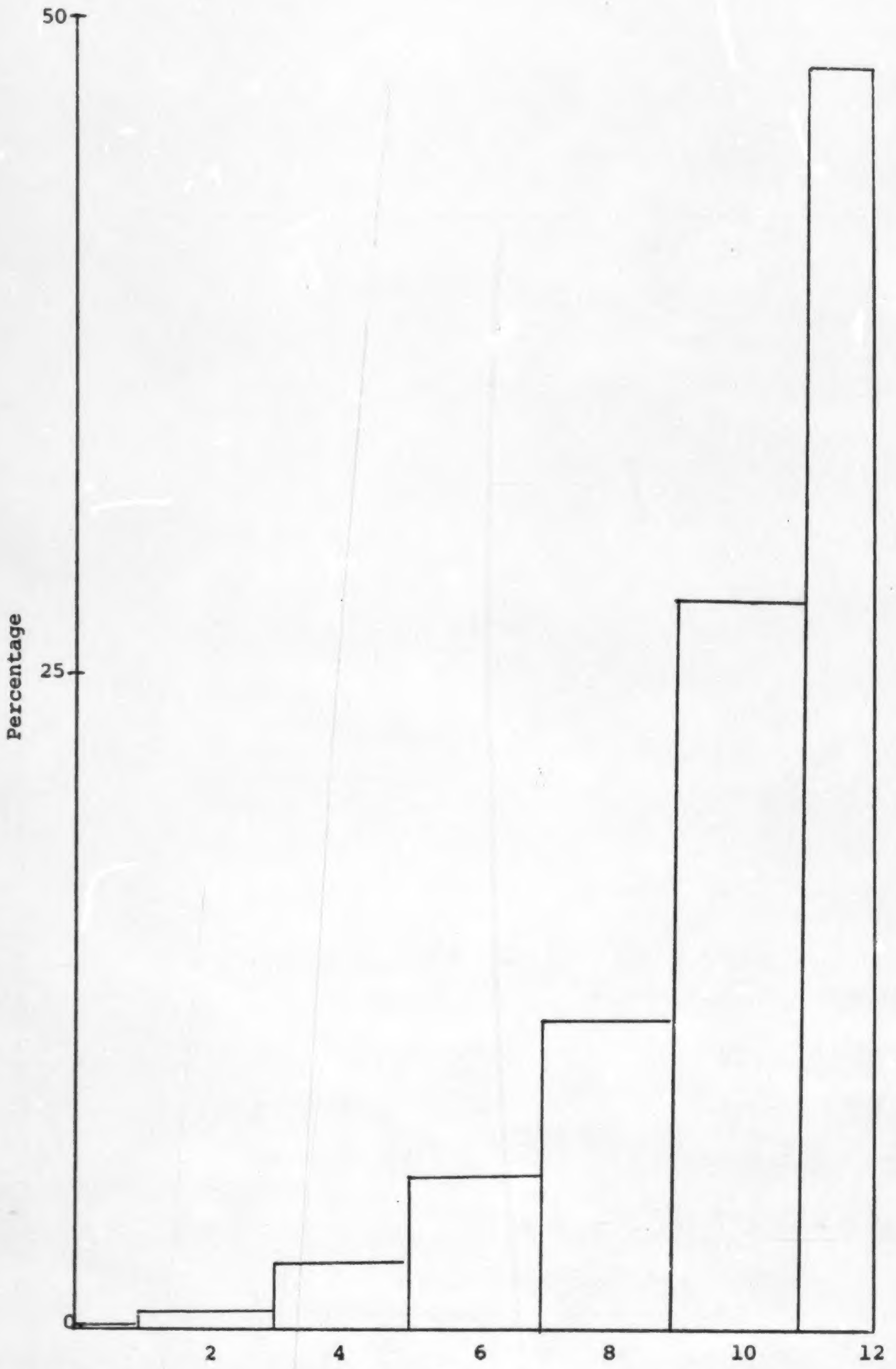


Figure B-21

Raw Scores of Kindergarten Gumpgookies Purposive Behavior Subtest

Random Sample of Population

N=2688

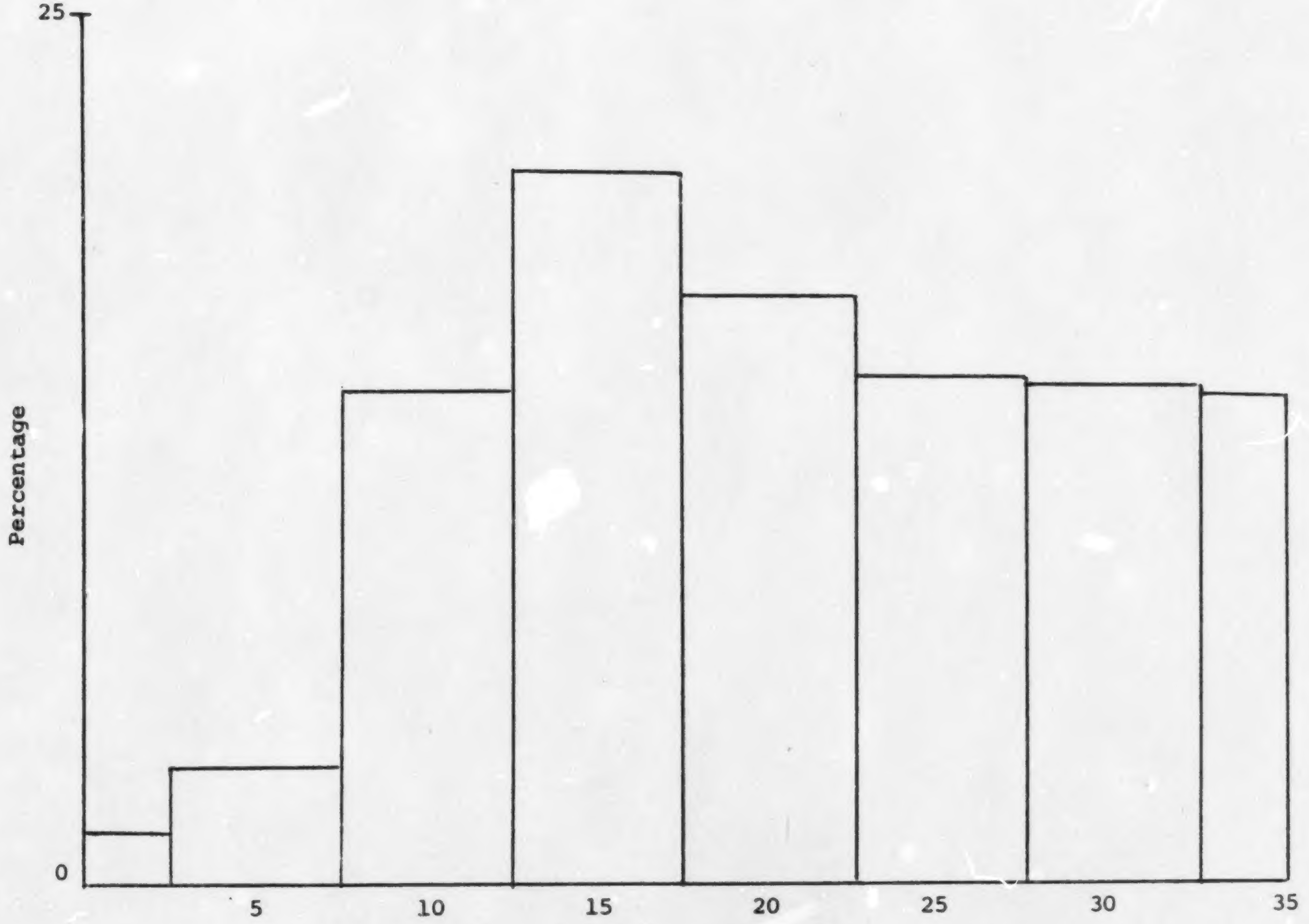


Figure B-22

Raw Scores of First Grade MAT Word Knowledge Subtest

Random Sample of Population
N=3061

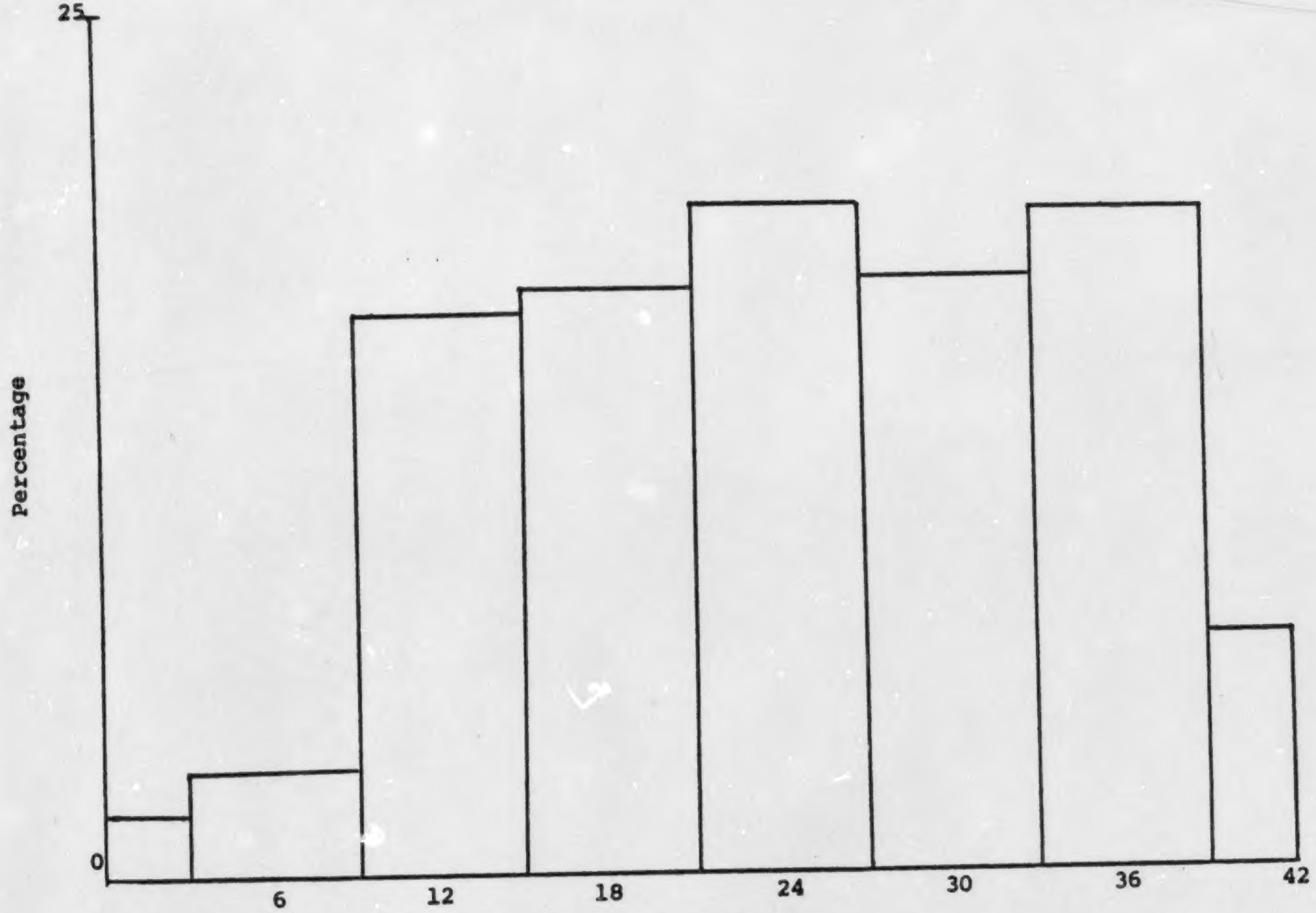


Figure B-23

Raw Scores of First Grade MAT Word Analysis Subtest

Random Sample of Population
N=3061

APPENDIX C:

RELIABILITIES, VALIDITIES AND COMPARATIVE DATA

APPENDIX C

RELIABILITIES, VALIDITIES AND COMPARATIVE DATA

1.0 INTRODUCTION

The batteries of instruments which have been administered to the Cohort III pupils consist of tests which have been carefully developed, tested, and used with many groups of children. It is, however, mandatory that we ascertain if the instruments are measuring our Cohort III pupils similarly to the way they measure other groups of pupils. We have, therefore, explored the reliability and validity characteristics of the instruments with this third Cohort. We have also compared Cohort III with the publisher's norming samples on several test characteristics to further elucidate the extent to which they differ.

2.0 RELIABILITIES

The overall reliability coefficients for the random sample of the population (hereafter referred to as the population) and for the ANCOVA analytic sample[1] appear in Tables C-1 and C-2. With the exception of the PPVT coefficient, which is a test-retest coefficient, they are all coefficients of internal consistency. These coefficients describe the extent to which the items in the subtest measure the same thing. We shall first examine the population coefficients and then compare these with the sample coefficients.

At the kindergarten level the coefficients for the Fall and Spring WRAT Total and Reading subtests are the highest (.91; .92). The three MAT subtest coefficients are all in the mid eighties (Numbers, .87; Listening for Sounds, .86; Reading, .84) as are the two WRAT Spelling subtests (.83). For the two WRATs and the MAT at the kindergarten level, the only low coefficients are the two WRAT Arithmetic subtests (Fall, .63; Spring, .72). The pretest-posttest coefficient for the PPVT is .67. One of the kindergarten affective tests, the Gumpgookies Total has a fairly high coefficient (.90); two have rather low coefficients (Gumpgookies Instrumental Activities, .58; Locus of Control-positive, .51) and one

---

[1] The ANCOVA analytic sample and the random sample of the population are defined in detail in Appendix A.

Table C-1

RELIABILITY COEFFICIENTS FOR A RANDOM SAMPLE OF THE POPULATION PRESENT FOR TESTING IN FALL 1971

Testing Time	Test	No. of Items	$\bar{X}$	S.D.	No. of Pupils	Rel.[a] Coeff.	SE of Measurement
Fall 1971	PEABODY PICTURE VOCABULARY TEST	75[b]	43.68	12.38	2,682	.67[c]	7.11
	WIDE RANGE ACHIEVEMENT TEST		28.22	12.82	6,401	.91	3.85
	Spelling	27[d,e]	4.90	4.10	6,401	.83	1.69
	Reading	38[d,e]	13.52	6.35	6,401	.91	1.90
	Arithmetic	17[d,e]	9.80	4.57	6,401	.63	2.78
Spring 1972	WIDE RANGE ACHIEVEMENT TEST		47.16	12.88	2,680	.92	3.64
	Spelling	27[d,e]	10.75	4.82	2,680	.83	1.98
	Reading	38[d,e]	21.48	6.46	2,680	.92	1.83
	Arithmetic	17[d,e]	14.92	4.05	2,680	.72	2.14
	METROPOLITAN ACHIEVEMENT TEST						
	Primer - Form F						
	Listening for Sounds	39	19.44	7.12	2,996	.86	2.66
	Reading	28	14.22	5.20	2,996	.84	2.08
	Numbers	34	12.96	6.00	2,996	.87	2.16
	PEABODY PICTURE VOCABULARY TEST	75[b]	51.03	10.97	2,682	.67[c]	6.30
	GUMPGOOKIES	60[d]	47.63	8.99	2,677	.90	2.84
	Instrumental Activities	12	9.29	1.93	2,677	.58	1.25
	LOCUS OF CONTROL						
	Internal Positive	11	5.62	2.08	2,695	.51	1.46
	Internal Negative	9	4.54	1.69	2,695	.33	1.39
Spring 1973	METROPOLITAN ACHIEVEMENT TEST						
	Primary 1 - Form F						
	Word Analysis	40	24.70	9.52	2,938	.93	2.52
	Reading	42	19.74	9.71	2,938	.92	2.75
	Reading Total (word knowledge plus reading)	77	41.07	17.10	2,938	.95	3.82
	Mathematics	62	30.82	12.91	2,938	.94	3.16

[a] Coefficients of internal consistency, except where noted.

[b] The first 75 items only were scored.

[c] Test - retest, Fall 1971, Spring 1972.

[d] Only selected items from the original test were administered.

[e] One of the items in the Spelling Subtest is worth two points, therefore, the highest score possible is 28. One of the items in the Reading Subtest is worth two points, therefore, the highest score possible is 39. One of the items in the Mathematics Subtest is worth eight points, therefore, the highest score possible is 24.

7

Table C-2

RELIABILITY COEFFICIENTS FOR ANCOVA ANALYTIC SAMPLE

Testing Time	Test	No. of Items	$\overline{X}$	S.D.	Rel. Coeff.[a]	SE of Measurement
Fall 1971	PEABODY PICTURE VOCABULARY TEST	75[b]	44.46	12.41	.76[c]	6.08
	WIDE RANGE ACHIEVEMENT TEST	82	29.59	12.39	.91	3.78
	Spelling	27[d,e]	5.29	4.19	.83	1.74
	Reading	38[d,e]	13.88	6.17	.91	1.89
	Arithmetic	17[d,e]	10.42	4.36	.62	2.70
Spring 1972	WIDE RANGE ACHIEVEMENT TEST	82	49.07	12.20	.92	3.45
	Spelling	27[d,e]	11.12	4.68	.82	1.99
	Reading	38[d,e]	22.48	6.18	.91	1.85
	Arithmetic	17[d,e]	15.47	3.81	.72	2.02
	METROPOLITAN ACHIEVEMENT TEST					
	Primer - Form F					
	Listening for Sounds	39	19.84	7.25	.86	2.68
	Reading	28	14.76	5.17	.84	2.08
	Numbers	34	[illegible]	6.34	.88	2.17
	PEABODY PICTURE VOCABULARY TEST	75[b]	55.37	7.54	.76[c]	3.69
	GUMPGOOKIES	60[d]	48.37	8.35	.88	2.84
	Instrumental Activities	12	9.44	1.82	.59	1.17
	LOCUS OF CONTROL					
	Internal Positive	11	5.74	2.04	.49	1.46
	Internal Negative	9	4.59	1.64	.28	1.39
Spring 1973	METROPOLITAN ACHIEVEMENT TEST					
	Primary 1 - Form F					
	Word Analysis	40	25.78	9.52	.93	2.54
	Reading	42	21.10	10.18	.93	2.78
	Reading Total (word knowledge plus reading)	77	43.64	17.53	.96	3.68
	Mathematics	62	32.15	13.09	.94	3.12

Note: N = 2755

[a] Coefficients of internal consistency, except where noted.

[b] The first 75 items only were scored.

[c] Test - retest, Fall 1971, Spring 1972.

[d] Only selected items from the original test were administered.

[e] One of the items in the Spelling Subtest is worth two points, therefore, the highest score possible is 28.
One of the items in the Reading Subtest is worth two points, therefore, the highest score possible is 39.
One of the items in the Mathematics Subtest is worth eight points, therefore, the highest score possible is 24.

quite low (Locus of Control-negative, .33). These low coefficients may in part be an artifact of test length (Gumpgookies Instrumental Activities, 12 items; Locus of Control-positive, 11 items; Locus of Control-negative, 9 items). However, it is likely that the Locus of Control-negative subtest is not internally consistent, but rather is measuring a diversity of information. The other two affective subtests correlate relatively well with the achievement subtests suggesting that even with their low homogeneity coefficients they are performing consistently.

The reliability coefficients for our ANCOVA analytic sample are very similar to those of our population with only two exceptions. The PPVT coefficient for the ANCOVA analytic sample is higher (.76) and the coefficient for the Locus of Control-negative is lower (.28).

The coefficients for the MAT subtests administered in the first grade range from a low of .92 for Reading to a high of .95 for Total Reading. The ANCOVA analytic sample coefficients are almost identical to these population coefficients, ranging from .93 for Reading to .96 for Total Reading.

Generally, the various reading subtests are internally consistent across both years. The arithmetic subtests become more homogeneous overtime. These findings would suggest that a diversity of information is being tapped by the WRAT Arithmetic subtests, and that different children are responding differentially to it. The other achievement tests seem to measure a relatively homogeneous set of information. A further discussion of the reliabilities, especially as they occur within different age groups, will be presented below (section 4) in a comparison of our ANCOVA analytic sample and the norming samples.

An important question for our ANCOVA strategy arises from a study of the test reliabilies. Are the psychometric instruments sufficiently reliable to be used as covariates or do they need to be corrected for falibility? We can see from the analytic design (Chapter II) that the Fall and Spring kindergarten WRATs are both necessary covariates if we wish to examine the first grade achievement tests in a longitudinal framework. The information presented in this section indicates that the total scores for both the Fall and Spring WRATs are very reliable (.91 and .92) for these covariates. Falibility corrections will not be necessary.

3.0 VALIDITIES

The subtests in any one of the four achievement batteries (Fall kindergarten WRAT; Spring kindergarten WRAT; Spring kindergarten MAT; Spring first grade MAT) are assumed to measure different aspects of a child's academic performance. Some of the subtests across time points are identical (e.g., Fall and Spring WRAT Spelling); others are similar (e.g., kindergarten MAT Reading and first grade MAT Reading). We would expect that under ideal conditions the subtests which are identical or similar would correlate more highly with each other than the subtests which are dissimilar or which measure different domains. This type of logic should extend to the affective tests as well. We should expect that their correlations with the achievement tests would be lower; however, if they are measuring affective traits which effect the academic domains differentially, this should be reflected in the relative size of the correlation coefficients.

There are a number of circumstances which may disrupt the pattern however. The subtests may not be as similar as they appear to be; different types of learning experiences for different types of children may disrupt patterns; and in our particular sample, the removal of zero score data from all but the Fall WRAT subtests may effect correlations differentially depending on which tests are being examined and which tests were most influenced by this action. Additionally, method variance may intrude; that is, it may be that the children perform more similarly on MAT subtests than they do across a MAT and a WRAT subtest because of within battery similarity of presentation, e.g., a group versus an individually administered test. It might also be that for a given testing cycle the children are more responsive to test taking than for another testing cycle. With as large and diverse a sample as we have, however, this latter circumstance is not likely to occur. Even with all these potential constraints, however, we should be able to observe a general pattern across the many subtests. We shall first examine the relationships among the achievement tests and then explore their relationship with the affective tests. We shall explore these relationships for our total analytic sample of 2755 children. The correlations are displayed in Table C-3.

The WRAT Spelling subtest appears to be fairly consistent in its

Table C-3

INTERTEST CORRELATION MATRIX FOR ANCOVA ANALYTIC SAMPLE

	FALL PPVT	FALL WRAT	WRAT SPELL	WRAT READ	WRAT ARITH	SPRING WRAT	WRAT SPELL	WRAT READ	WRAT ARITH	MAT SOUNDS	MAT READ	MAT NUMB	SPRING PPVT	GUMP TOT	GUMP ACT	LOCUS INT POS	LOCUS INT NEG	MAT WORD AN	MAT READ	READ TOT	MAT MATH	TRANSFORMED GUMP ACT	TRANSFORMED GUMP TOTAL	TRANSFORMED MAT WORD AN
FALL WRAT	.53	-																						
WRAT SPELL	.42	.78																						
WRAT READ	.45	.89	.52																					
WRAT ARITH	.45	.83	.52	.62																				
SPRING WRAT	.45	.73	.59	.63	.62																			
WRAT SPELL	.31	.54	.57	.41	.40	.78																		
WRAT READ	.38	.64	.43	.63	.53	.88	.49																	
WRAT ARITH	.42	.63	.49	.50	.62	.81	.46	.62																
MAT SOUNDS	.37	.54	.44	.47	.46	.63	.45	.58	.53															
MAT READ	.33	.57	.41	.53	.48	.71	.45	.72	.54	.66														
MAT NUMB	.40	.59	.50	.46	.54	.63	.46	.51	.63	.57	.55													
SPRING PPVT	.76	.48	.40	.41	.41	.43	.29	.38	.42	.37	.32	.39												
GUMP	.26	.25	.22	.20	.23	.28	.22	.22	.26	.23	.22	.25	.33											
GUMP ACT	.20	.19	.17	.15	.17	.22	.18	.17	.21	.19	.18	.17	.26	.83										
LOCUS INT POS	.37	.33	.28	.26	.31	.33	.22	.28	.32	.27	.23	.35	.37	.22	.17									
LOCUS INT NEG	.17	.14	.12	.09	.14	.16	.10	.15	.16	.14	.12	.13	.19	.10	.08	.24								
MAT WORD AN	.47	.61	.49	.52	.51	.71	.50	.64	.60	.57	.57	.53	.44	.24	.20	.29	.13							
MAT READ	.39	.57	.47	.50	.46	.64	.48	.57	.53	.57	.56	.54	.36	.21	.17	.28	.12	.75						
MAT READ TOT	.43	.62	.50	.54	.50	.69	.51	.63	.57	.59	.59	.56	.41	.23	.19	.29	.13	.85	.95					
MAT MATH	.46	.59	.50	.46	.55	.64	.48	.51	.62	.47	.48	.62	.43	.24	.19	.32	.14	.69	.66	.71				
GUMP ACT	.20	.19	.17	.14	.17	.21	.17	.16	.20	.18	.17	.17	.26	.78	.97	.16	.08	.20	.16	.18	.19			
GUMP TOTAL	.29	.27	.24	.21	.25	.29	.23	.22	.28	.25	.23	.26	.35	.93	.76	.24	.11	.26	.23	.25	.26	.75		
MAT WORD AN	.47	.61	.50	.53	.51	.71	.51	.64	.60	.58	.59	.54	.44	.24	.20	.29	.13	.99	.77	.86	.70	.20	.26	

Note: N = 2755

performance. It correlates more highly with itself (.57) across time than it does with either the Reading or Arithmetic subtests administered either at the same time point or across time.

The WRAT Reading and Arithmetic subtests hold up fairly well although it appears that they are not as differentiated as is the Spelling. Each does correlate higher with itself across time [illegible] with the others at the same point in time; i.e., the Fall and Spring WRAT Reading correlate, .63. the Fall and Spring WRAT Arithmetic correlate, .62; however, the Fall WRAT Reading and Arithmetic also correlate .62 as do the Spring WRAT Reading and Arithmetic.

This does not necessarily suggest that the Arithmetic and Reading subtests are measuring the same thing. It may suggest, however, that comparable levels of skill development are being tapped in each of these two subtests, i.e., a child who has a given level of skill development in the arithmetic areas being measured is likely to possess approximately the same level of skill development in the concurrent reading areas.

The Primer and Primary I levels of the MAT, administered at the end of kindergarten and first grade respectively, have different, although somewhat similar, subtests. We might expect that the kindergarten Listening for Sounds and the first grade Word Analysis subtests would require similar types of skills, although at different developmental levels. Similarly, the kindergarten Reading subtest and the first grade Reading and Total Reading subtests may all be measuring a diversity of skills, although at different levels. We would also expect that the arithmetic subtests would be more correlated with each other than with the reading subtests.

Turning to the MAT subtest correlations, we find that they do not conform to our ideal pattern although they do approximate it to varying degrees. At the kindergarten level, Listening for Sounds and Reading subtests have a higher correlation with each other (.66) than they do with their similar subtests at the first grade level (.56 to .59); they correlate as highly with the kindergarten Numbers test (.55; .57) as they do with the first grade reading subtests (.56 to .59); however, their correlations with the first grade Mathematics are lower (.47). This is the same type of pattern which we noted for the WRAT Arithmetic and Reading subtests.

At the first grade level the Total Reading score includes the Reading subtest. Since the two scores tend to perform in a similar fashion, we shall examine only the Total Reading along with the Word Analysis subtest. Here, although the correlations are higher, the same type of pattern emerges. The two subtests have a higher correlation with each other (.85) than they do with similar MAT subtests at the kindergarten level (.57, .59). They also correlate higher with the first grade Mathematics subtest (.69, .71) than with the kindergarten reading subtests. Their lowest correlations are with the kindergarten Numbers subtest (.53, .56).

The kindergarten Numbers subtest has a relatively low correlation with the first grade Mathematics subtest (.62); however, it is higher than its correlations with any of the MAT reading subtests at either Kindergarten or first grade level (.53 to .57).

The first grade Mathematics subtest conforms according to the characteristic pattern. Its correlations with the first grade reading subtests (.69, .71) are higher than its correlations with the kindergarten Numbers subtest (.62), which in turn is higher than its correlations with the kindergarten reading subtests (.47).

Turning now to a comparison of the two administrations of the WRAT and the two MATs, we note that in general, as we would expect the Spring WRAT subtests scores correlate higher than the Fall WRAT subtest scores with both the Spring kindergarten MAT and the first grade MAT. However, what might not be expected is that the first grade MAT scores tend to correlate higher with both WRATs than do the kindergarten MAT scores.

The correlations among the two administrations of the PPVT (Fall and Spring), which is a vocabulary test, and the various WRAT subtests range from .31 to .47. Both the Fall and Spring PPVTs have as high or higher correlations with the arithmetic subtests and the first grade Word Analysis subtest as with any of the other reading subtests of the WRATs and MATs. The highest correlations are between the PPVT scores and the Fall WRAT and the first grade MAT scores (.36 to .47).

Four affective scores are available: Gumpgookies Total; Gumpgookies Instrumental Activities; Locus of Control-positive and Locus of Control-negative. Other than the correlation between the Gumpgookies Total and subtest scores, these tests do not correlate highly among themselves (.08

to .34), nor with the achievement domain (.09 to .37). There does appear, however, to be a pattern of relationship. In general, Locus of Control-positive correlates higher with the achievement tests (.22 to .37), followed by the Gumpgookies Total (.21 to .35) the Gumpgookies Instrumental Activities (.14 to .26), the Locus of Control-negative (.09 to .19). The highest correlations occur between Locus of Control-positive and the two PPVT scores (.37). Following this, Locus of Control-positive tends to correlate more highly with the arithmetic subtests than with the reading subtests.

Overall, in the achievement area, we have found that similar subtests tend to correlate more highly among themselves at a given point in time, followed by different domain subtests at the same point in time; than similar subtests across time and finally different domain subtests across time. There is also a general tendency for the first grade MAT scores to correlate higher with the Fall and Spring WRATs and PPVTs than do the kindergarten MAT scores. Locus of Control-positive appears to be performing at a better predictor of achievement scores than the other affective tests.

These findings suggest that although many of the correlations are not as high as desired and the pattern of relationship does not reflect the ideal, the subtests generally work within their respective domains and that there are some observable patterns occurring between the academic and affective domains.

4.0 NORMATIVE DESCRIPTION

We have now discussed the reliability and validity characteristics of the achievement and affective batteries as they pertain to this year's ANCOVA analytic sample and a random sample of the FT/NFT population. Because the pupils in this population and sample have been specifically selected for the FT treatment or for a comparison group, they are not expected to be similar to the groups with which these batteries were originally developed and normed. Therefore the tests may not be functioning with these groups of pupils as they are expected to function with children who are more similar to the original norm groups. The tests may be less reliable; they may even be measuring different skills or different levels of skill development.

For example, looking at the WRAT, we note that the Spelling subtest is made up of copying 18 marks, followed by printing at least two letters of the child's name correctly, and then taking a written spelling test. For low scoring children, we may be measuring a child's psychomotor development while for higher scoring children we may be measuring initial phonic or spelling skills. In the norming sample, six year olds attained a mean of 18.70 with a standard deviation of 5.70, suggesting that many of these children were able to copy marks and that the items which most differentiated them were the spelling words; that is that the Spelling subtest probably functioned for many of the children as a test of their early spelling performance. For our random sample of the population on the Fall WRAT Spelling test, the mean for six year olds was 5.88 with a standard deviation of 4.67, suggesting that most of these six year old pupils were having difficulty copying marks at kindergarten entry; therefore, the Fall test for this group is probably functioning as a test of a lower level skill. This is not to say that the test is inappropriate for these pupils; it is to point out that tests may be measuring different skills or different developmental levels of skills in our groups from those it measures in other groups. We shall now look at each of the achievement batteries in turn to try to ascertain if our pupils are performing similarly to the groups of pupils with whom the tests were developed.

WRAT. As mentioned above, each of the subtests of the WRAT was administered in a shortened version, with the more difficult items omitted. This resulted in highest possible scores of 28 in Spelling, 39 in Reading, and 24 in Arithmetic. These high scores represent grade levels of 2.2, 2.1, and 2.4 respectively. Since it is unlikely that many children in the tested groups would achieve such performance levels, the restriction in test scores is of relative unimportance when utilizing raw scores. However, if one wishes to compare the performance of these children with the performance of a norming sample, it becomes necessary to determine potential restrictions at the top of the range of scores. Table E-4 presents the number and percentage of children achieving within ten points of the highest possible score on each of the administered subtests. It also presents that data for the total group as additional information although total score characteristics are not discussed by the publisher.

There appears to be no need for any concern with the Fall adminis-

Table C-4

NUMBER OF PUPILS IN RANDOM SAMPLE OF POPULATION ATTAINING SCORES WITHIN TEN POINTS OF HIGHEST POSSIBLE SCORE ON WRAT SUBTESTS

Analytic Sample

	Spelling			Reading			Arithmetic			Total		
	Score	N	%	Score	N	%	Score	N	%	Score	N	%
Fall	19	11	.4	30	0	0	15	217	7.9	82	0	0
	20	1	0	31	1	0	16	158	5.7	83	0	0
	21	0	0	33	1	0	17	91	3.3	84	0	0
	22	0	0	34	2	.1	18	51	1.9	85	0	0
	23	0	0	36	1	.0	19	24	.9	86	0	0
	24	0	0	37	1	.0	20	11	.4	87	0	0
	25	0	0	39	3	.1	21	1	0	88	0	0
	26	0	0				22	1	0	89	0	0
	27	0	0				23	0	0	90	0	0
	28	0	0				24	0	.0			
Spring	19	58	2.1	30	28	1.0	15	317	11.5	82	3	.1
	20	57	2.1	31	23	.8	16	370	13.4	83	2	.1
	21	26	.9	32	20	.7	17	355	12.9	84	1	0
	22	15	.5	33	20	.7	18	234	8.5	85	2	.1
	23	12	.4	34	19	.7	19	171	6.2	86	3	.1
	24	8	.3	35	18	.7	20	146	5.3	87	1	0
	25	6	.2	36	18	.7	21	85	3.1	88	0	0
	26	5	.2	37	13	.5	22	54	2.0	89	0	0
	27	2	.1	38	19	.7	23	37	1.3	90	0	0
	28	0	0	39	0		24	35	1.3	91	0	0

POPULATION

	Spelling			Reading			Arithmetic			Total		
	Score	N	%	Score	N	%	Score	N	%	Score	N	%
Fall	19	8	.1	28	22	.3	15	418	6.5	82	0	0
	20	1	0	29	7	.1	16	322	5.0	83	0	0
	22	2	0	30	3	0	17	183	2.9	84	1	0
	24	1	0	31	3	0	18	102	1.6	85	0	0
	25	0	0	32	2	0	19	44	.7	86	0	0
	26	0	0	34	3	0	20	25	.4	87	0	0
	27	0	0	36	2	0	21	7	.1	88	0	0
	28	0	0	37	1	0	22	1	0	89	0	0
				38	2	0	23	4	.1	90	0	0
				39	4	.1	24	0	0	91	0	0
Spring	19	58	2.2	28	93	3.4	15	305	11.2	82	1	0
	20	40	1.5	29	55	2.0	16	334	12.3	83	4	.1
	21	17	.6	30	34	1.3	17	305	11.2	84	1	0
	22	20	.7	31	24	.9	18	221	8.1	86	3	.1
	23	7	.3	32	17	.6	19	169	6.2	87	2	.1
	24	5	.2	33	16	.6	20	122	4.5	88	2	.1
	25	4	.1	34	18	.7	21	72	2.7	90	1	0
	26	4	.1	35	17	.6	22	47	1.7	91	0	0
	27	4	.1	36	14	.5	23	24	.9			
	28	0	0	37	12	.4	24	23	.8			
				38	6	.2						
				39	14	.5						

tration of the Spelling subtest. Less than one percent of the scores might have been effected to any great extent by the Fall administration of the Reading and Arithmetic subtests or the Spring administration of the Spelling subtest. A somewhat larger percentage of scores may have restricted in the Spring Reading subtest. It is fairly reasonable to assume that a sizeable number of Spring Arithmetic subtest scores have been depressed. With one out of the six administered subtests showing potentially serious symptoms of topping, it appears to be reasonable to compare distributional characteristics of our pupils and of the norming sample group. Emphasis needs to be given to the point that some children would probably have achieved higher scores on the Spring Arithmetic subtest if it had been administered according to the publisher's directions.

The publisher provides means and standard deviation at six month age intervals, beginning at age five. We have presented Fall 1971 and Spring 1972 data for our ANCOVA analytic sample at corresponding six month age intervals. (See Table E-5.) Because of the small number of children in many of these age level groupings within our sample, and because the publisher provides no information below age five, we shall discuss only age groupings 5, 5 1/2, and 6.

For each of the three subtests, as age increases the subtest score also increases for the publisher's norming sample. This expected phenomenon also occurs in our analytic sample at both testing points, although perhaps not to as great an extent.

There are, however, some dramatic differences between the main performance levels of our groups and the norming sample, with our groups scoring considerably lower. In general, our Spring scores are closer than our Fall scores to the norms. At the individual subtest level, the greatest divergence of scores occurs with the Spelling subtest. Comparing age levels, we note that our 5 and 5 1/2 year olds at the Spring testing point appear to be performing at a level higher than the 5 and 5 1/2 year olds in the norming sample on the Reading subtest. These same groups appear to be performing at a slightly higher level than the norming sample on the Arithmetic subtest. Generally, our 6 year olds perform at a lower level than do the norming sample 6 year olds.

The reliability and standard errors of measurement of the WRAT sub-

Table C-5

COMPARATIVE DATA ON WRAT (LEVEL I): MEANS AND STANDARD DEVIATIONS FOR PUBLISHER AND FOR ANCOVA ANALYTIC SAMPLE FOR AGES FIVE TO TEN

PUBLISHER

	Spelling			Reading			Arithmetic		
Age [a]	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N
5	13.87	5.26	191	16.96	7.70	223	12.74	4.41	191
5½	15.35	5.57	304	19.37	7.75	310	13.51	4.49	304
6	18.70	5.70	334	27.08	10.43	340	17.14	4.70	331
6½	23.50	5.89	304	32.25	10.72	327	19.32	4.74	301
7	27.03	6.13	339	39.85	12.30	358	22.16	4.62	337
7½	31.15	6.30	357	46.83	12.40	389	24.97	4.68	357
8	33.56	6.47	332	51.48	12.68	367	27.03	4.51	332
8½	35.28	6.66	335	54.17	13.04	357	28.50	4.42	334
9	37.56	6.91	321	57.40	13.07	341	29.95	4.43	321
9½	46.07	7.35	296	61.16	13.37	328	21.93	4.80	296
10	42.15	7.94	308	63.18	13.44	330	33.72	5.77	306
ANCOVA ANALYTIC SAMPLE: FALL 1971									
3½	2.75	2.99	4	12.00	7.70	4	8.00	4.08	4
4	1.00	1.73	3	8.67	.58	3	4.00	2.65	3
			196						
4½	3.63	3.53	196	12.62	6.19	196	9.05	3.85	196
5	4.50	3.84	1209	13.31	5.95	1209	9.70	4.22	1209
5½	5.98	4.17	1151	14.41	6.25	1151	11.05	4.43	1151
6	8.24	4.96	165	16.10	6.43	165	12.44	3.94	165
6½	7.18	4.66	22	13.55	5.95	22	11.09	4.63	22
7	1.67	.58	3	15.67	7.77	3	11.00	6.08	3
7½	6.00	-	1	11.00	-	1	17.00	-	1
9½	0.00	-	1	10.00	-	1	2.00	-	1
ANCOVA ANALYTIC SAMPLE: SPRING 1972									
4	17.00	-	1	19.00	-	1	15.00	-	1
4½	7.83	4.62	6	17.17	7.28	6	14.00	2.45	6
5	9.44	4.70	97	20.04	5.91	97	14.04	3.58	97
5½	10.61	4.60	1099	21.50	5.82	1099	15.01	3.84	1099
6	11.47	4.64	1232	21.85	5.84	1232	15.68	3.77	1232
6½	12.18	4.76	288	23.09	5.70	288	16.73	3.64	288
7	11.69	5.33	26	22.15	4.10	26	16.39	3.23	26
7½	11.67	2.31	3	21.33	3.79	3	15.00	2.65	3
8	11.50	4.95	2	23.00	2.83	2	16.00	1.41	2
10	5.00	-	1	22.00	-	1	11.00	-	1

[a] N = 0 for Age groups missing from the ANCOVA Analytic Sample

tests are presented in Table C-6. The publisher's norming coefficients are split-half, based on items which have been ordered according to difficulty level. Ours are coefficient alphas which would probably be slightly lower than the type used by the publisher. Our coefficients are, however, in most cases, substantially lower, suggesting that the tests are not as reliable when used with our groups. Our Spring test coefficients are higher than our Fall coefficients. This, in conjunction with the finding that our Spring means are more similar to the publisher's than our Fall means, suggests that our sample's performance was more similar to the norming sample performance on the Spring WRAT than it was on the Fall WRAT. However even in the Spring there are differences between the norming sample and ours.

PPVT. For both the Fall and Spring administrations of the PPVT, only the first seventy five items were scored. This decision is the result of a preliminary examination of these data for the random sample of the population. Less than one percent of these pupils answered any single item correctly beyond item 75. Among the items answered correctly beyond this point were words such as vitreous, nape, concentric, bereavement, and hoisting. The reliability of the information obtained from such data would be open to question. Table C-7 presentsthe number of pupils who would have attained scores of 60 or higher on each of the two administrations of the PPVT if the complete tests (150 items) had been scored. These data demonstrate that the cut off point of 75 items restricted scores of very few pupils. We therefore have chosen to utilize information on only the first 75 items, believing that this cut off point neither restricts the potential range of test scores nor introduces data of questionable quality.

The publisher's data and our comparative ANCOVA analytic sample data are displayed in Table C-8. These data are provided by the publisher at varying age intervals. Our data are given for similar age groupings. We shall be primarily concerned with only the two age groups which contain the most children in our ANCOVA analytic sample (4 yrs. 9 mos. to 5 yrs. 5 mos.; 5 yrs. 6 mos. to 6 yrs. 5 mos.). At the Fall testing point our ANCOVA analytic sample's performance was both considerably lower and more variable for both age groups. In the Spring, our ANCOVA analytic sample's performance at a given age level was higher than it had been in the Fall.

Table C-6

COMPARATIVE DATA ON WRAT (LEVEL I): RELIABILITIES FOR PUBLISHER'S NORMING SAMPLE AND ANCOVA ANALYTIC SAMPLE

		Publisher's Norming Sample			ANCOVA Analytic Sample					
					Fall 1971			Spring 1972		
	Age	N	r	SEM	N	r	SEM	N	r	SEM
WRAT Spelling	5	200	.97	.86	2356	.82	1.73	1193	.82	1.98
	6	200	.95	1.04	186	.85	1.91	1519	.82	1.99
WRAT Reading	5	200	.98	1.10	2356	.90	1.92	1193	.92	1.78
	6	200	.99	1.36	186	.91	1.92	1519	.92	1.74
WRAT Arithmetic	5	200	.97	.88	2356	.61	2.73	1193	.70	2.10
	6	200	.96	.98	186	.69	2.25	1519	.73	1.94

Table C-7

NUMBER OF PUPILS IN RANDOM SAMPLE OF POPULATION ACHIEVING A SCORE OF 60 OR MORE ON THE FALL 1971 AND SPRING 1972 ADMINISTRATION OF THE 150-ITEM PPVT (FORM A)

	FALL		SPRING	
Score	N	%	N	%
Below 60	5885	93.8%	2210	81.6%
61	84	1.3	68	2.5
62	77	1.2	82	3.0
63	73	1.2	66	2.4
64	39	.6	62	2.3
65	27	.4	36	1.3
66	19	.3	37	1.4
67	20	.3	29	1.1
68	16	.1	19	.7
69	9	.1	19	.7
70	6	.0	12	.4
71	3	.0	13	.5
72	2	.0	9	.3
73	2	.0	9	.3
74	3	.0	8	.3
75	1	.0	7	.3
76	1	.0	6	.2
77	2	.0	2	.1
78	1	.0	2	.1
79	2	.0	2	.1
80	0	.0	3	.0
81	0	.0	3	.1
82	0	.0	0	.1
83	0	.0	0	.0
84	1	.0	0	.0
85-	0	.0	1	.1
86	0	.0	0	.0
87	0	.0	0	.0
88	1	.0	0	.0
89	1	.0	1	.1
	6275	100.0%	2707	100.0%

Table C-8

COMPARATIVE DATA ON PPVT (FORM A): MEANS AND STANDARD DEVIATIONS FOR PUBLISHER AND FOR ANCOVA ANALYTIC SAMPLE FOR AGES 3^9 TO 10^5

	Publisher's Norming Sample		ANCOVA Analytic Sample					
			Fall 1971			Spring 1972		
Age [a]	$\overline{X}$	S.D.	N	$\overline{X}$	SD	N	$\overline{X}$	SD
3^9 to 4^2	42.08	10.34	6	31.33	11.73	-	-	-
4^3 to 4^8	45.58	8.64	3	39.00	3.61	6	48.17	5.98
4^9 to 5^5	50.22	8.17	1403	42.52	12.22	98	49.19	10.12
5^6 to 6^5	55.37	7.52	1316	46.62	12.26	2331	52.65	9.37
6^6 to 7^5	60.82	7.77	25	42.96	11.78	314	57.51	8.07
7^6 to 8^5	65.81	8.60	1	62.00	-	5	50.40	8.79
9^6 to 10^5	75.49	9.64	1	33.00	-	1	51.00	-
		TOTAL	2755	44.46	12.41	2755	53.06	9.40

[a] N = 0 for age groups missing for ANCOVA analytic sample.

For the younger age group our sample PPVT mean was similar to that of the norming sample, but the variability was greater in our sample. For the older age group our sample's scores were both lower and more variable than the norming sample's scores.

Our sample performed more similarly to the norming sample in the Spring than it did in the Fall; however, there are differences between the norming sample's performance and our sample's performance at both test points.

The publisher provides alternate form reliability for these data. Our test-retest coefficient (pretest-posttest) was .76 for the ANCOVA analytic sample. Unfortunately these coefficients are not directly comparable.

MAT. The MAT, Primer Level was administered in the Spring of 1972 as the children in the analytic sample were completing kindergarten. In the Spring of 1973 our sample was administered the Primary I level, Form F of the MAT. The publisher provides norming sample information for these levels at grades k.7 and 1.7 respectively. The publisher's and ANCOVA analytic sample data are displayed in Table C-9.

At the kindergarten level the ANCOVA analytic sample's performance was quite similar to that of the publisher's norming sample although the reliabilities were not as high.

At the first grade level, there were some disparities in performance levels between the two groups. However, the reliability coefficients at this time are nearly identical. The two groups performed similarly on the Word Analysis subtest. On the Reading and Total Reading subtests differences appear to emerge. While the means are quite similar, the medians are not. The norming sample's median is numerically slightly higher than its mean while ours is considerably lower. This suggests that we may have had a few more high scorers in our ANCOVA analytic sample than did the norming sample.

Table C-9

COMPARATIVE DATA ON MAT (FORM F): MEANS, MEDIANS, AND STANDARD DEVIATIONS FOR PUBLISHER'S NORMING SAMPLE[a] AND ANCOVA ANALYTIC SAMPLE[b]

		Publisher's Norming Sample				ANCOVA Analytic Sample			
Level	Subtest	$\bar{X}$	Med	S.D.	r^c	$\bar{X}$	Med	S.D.	r^c
Primer	Listening for Sounds	19.3	19.1	8.52	.92	19.84	19.21	7.25	.86
	Reading	13.0	12.9	6.91	.92	14.76	14.61	5.17	.84
	Numbers	12.8	13.7	7.43	.90	13.53	11.87	6.34	.88
Primary I	Word Analysis	27.3	26.2	9.09	.92	25.78	26.32	9.53	.93
	Reading	19.8	21.9	9.81	.93	21.10	17.75	10.18	.93
	Total Reading	43.4	45.2	16.84	.96	43.64	39.17	17.53	.96
	Mathematics	39.6	38.1	13.11	.95	32.15	31.56	13.09	.94

[a] Primer administered at K 7; Primary 1 at 1.7.

[b] Primer administered in Spring 1972 (kindergarten); Primary 1 in Spring 1973 (first grade).

[c] Publisher's coefficients are split-half; ANCOVA analytic sample coefficients are coefficient alphas.